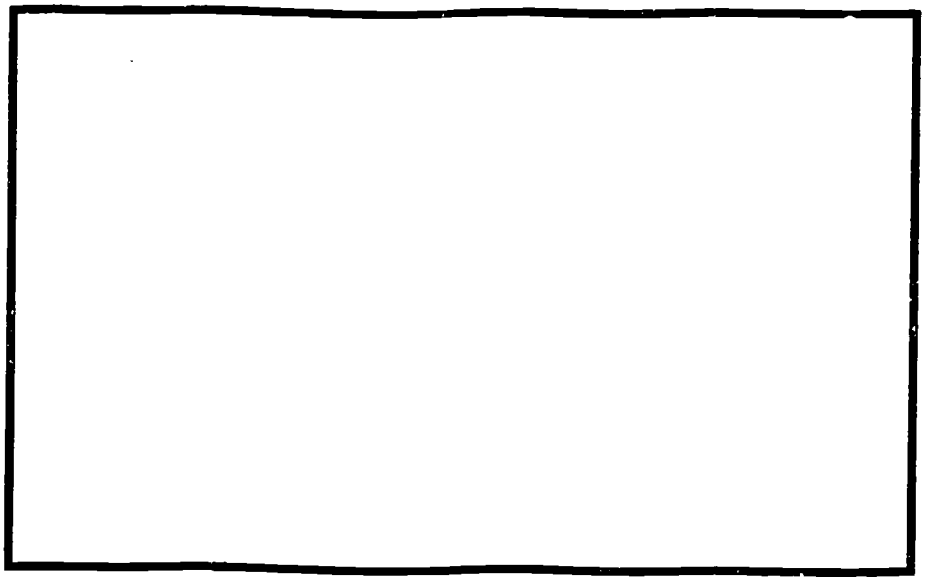
DOCUMENT RESUME

ED 084 882 EM 011 706

AUTHOR	Judd, Wilson A.; And Others
TITLE	Development and Formative Evaluation of a Five Module Computer-Managed Instructional System for Educational Psychology. Technical Report Number Nineteen.
INSTITUTION	Texas Univ., Austin. Computer-Assisted Instruction Lab.
SPONS AGENCY	National Science Foundation, Washington, D.C.
REPORT NO	TR-19
PUB DATE	Apr 73
NOTE	85p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Academic Performance; Behavioral Objectives; *Computer Assisted Instruction; Course Evaluation; Criterion Referenced Tests; Curriculum Development; *Educational Psychology; Formative Evaluation; Higher Education; Individualized Programs; Instructional Design; Instructional Materials; Program Costs; Teacher Education Curriculum
IDENTIFIERS	*Computer Managed Instruction

ABSTRACT

This report describes the design, development, and formative evaluation of a computer-managed instruction (CMI) system in a university setting. The system, implemented on an IBM 1500 Instructional System, is designed to serve an undergraduate educational psychology course for teacher trainees. Two broad goals are defined for the project: satisfactory post-instructional student performance; and positive affective outcomes with regard to the subject matter and the instruction/evaluation system. Subject matter is structured into five modules: 1) Computers in Education; 2) Classroom Management; 3) Tests and Measurement; 4) Cultural Differences; and 5) Statistics. The introduction, behavioral objectives and sample test items for each module are presented in appendices to the report. Only the data regarding performance measures are presented in this report. Evaluation of the affective aspects of the system will be discussed in subsequent reports. Based on project trials and initial use, modules and evaluation instruments have been revised. (JB)



THE UNIVERSITY OF TEXAS AT AUSTIN

Computer Assisted Instruction Laboratory

AUSTIN

DEVELOPMENT AND FORMATIVE EVALUATION OF A FIVE MODULE COMPUTER-MANAGED INSTRUCTIONAL SYSTEM FOR EDUCATIONAL PSYCHOLOGY

Technical Report No. 19

Wilson A. Judd, Harold F. O'Neil, Jr.,
Donald D. Rogers, & Frank C. Richardson

With the assistance of:

Paul N. Dixon
Jody L. Fitzpatrick
Ronald W. Mitchell
Brenda Rutherford
Douglas Saddler
W. Paul Scott
Claire E. Weinstein

April 1973

(First Draft December 1972)

Supported By:

THE NATIONAL SCIENCE FOUNDATION
Grant GJ 509 X

Computer-Assisted Instruction Laboratory
The University of Texas at Austin
Austin, Texas 78712

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Until fairly recently, major interest in the instructional use of computers has centered around various forms of computer-assisted instruction (CAI), all of which involve intensive interaction between the student and computer. Due to the amount of effort required to specify the details of this interaction adequately, the design and implementation of such programs is an expensive and time-consuming process. Consequently, most of the CAI programs currently available are fairly short and pertain to only a very small portion of a complete course. As a result, one often finds short segments of CAI embedded in a course which is otherwise taught in a traditional, non-individualized way. Furthermore, since they can bear only a small portion of the total instructional load, such short programs usually represent additional expense rather than approaching justification as a replacement cost.

An alternative approach is to make use of the computer's information processing and memory capabilities to manage the individualization of a complete course or the major portion of a complete course. Under a computer-managed instruction (CMI) system, the instruction itself is not presented by computer. Rather, the instruction is presented by means of conventional, less expensive media. While the most commonly used medium is the printed page, slide presentations, audio tapes and films could easily be incorporated. The student proceeds to study at his own pace while the computer is used to monitor and, to some extent, direct his progress through the instructional materials. This is done by means of testing at frequent points in the program to diagnose the student's strengths and weaknesses, providing prescriptions for remedial work, and scheduling students' use of the available instructional resources. Short segments of CAI might, of course, be among the resources available for assignment by the CMI system. While similar programs of individualized instruction have been conducted without recourse to a computer, the magnitude of the management activity required becomes overwhelming when large numbers of students are involved.

The purpose of this report is to describe the design, development, and formative evaluation of a CMI system in a university setting. The system was designed to serve an undergraduate educational psychology course for teacher trainees at The University of Texas at Austin. It was developed by The University of Texas Computer-Assisted Instruction Laboratory and implemented on an IBM 1500 Instructional System (IBM, 1967). The major source of funding for the project was the National Science Foundation.

Two broad goals were defined for the complete CMI project: satisfactory post-instructional student performance, and positive affective outcomes with regard to the subject matter and to the instruction/evaluation system. Only the data regarding performance measures will be presented in this report. Evaluation of the affective aspects of the system as measured by expressed student curiosity about the subject matter, reported anxiety during testing, and responses to an evaluation questionnaire, will be discussed in subsequent reports.

One major purpose for which the development of this system was funded was to provide a vehicle for educational research. While research has been and continues to be conducted within the context of the system, the research itself will not be discussed here but will be presented in detail in reports pertaining the specific projects (e.g., Richardson, O'Neil, Grant, & Judd, 1973). The environment in which this research is taking place and the factors which were instrumental in creating this environment will be described here.

INSTRUCTIONAL DEVELOPMENT

The student performance data to be reported were derived from student scores on criterion-referenced evaluation instruments designed to measure achievement of the instruction's objectives. Since scores on instruments of this type do not lend themselves to analysis by the traditional statistical measures of validity and reliability, the meaning of these scores may be seriously questioned unless there is a strong reason to believe that the procedures used for designing the instruction and developing the evaluation instruments were adequate (e.g., Glaser, 1963; Livingston, 1972; Kriewall, 1972; and Edmonston, Randall, & Oakland, 1972). For this reason, the characteristics of the product development group and the instructional design procedures followed will be discussed in detail.

Product Development Group

The instructional product development group consisted of the two project directors and seven graduate students, divided into two-person teams. Typically each team contained a content matter specialist and an instructional design specialist and each team was given the primary responsibility for developing one module. Because of a shortage of qualified personnel, a few of the students served simultaneously on two teams. The project directors provided the management required for the coordination of the teams. This included monitoring the work of all teams to insure that a standard of quality was maintained, providing the expertise necessary to solve problems beyond the competencies of the students, supervising the allocation of resources such as secretarial time, and providing liaison with other faculty members involved. Although the basic production unit was the two-man team, occasionally all teams met as a group to report progress and to have their work evaluated. However, most inter-team communication was mediated by the project directors. A schematic diagram of the management/communications network is presented in Figure 1.

PD--Project Directors
I--Instructional Design Specialist
C--Content Matter Specialist
⟷ --Indicates strong interaction
⇠⇢ --Indicates weak interaction
*--The same person filled both roles on Team Four

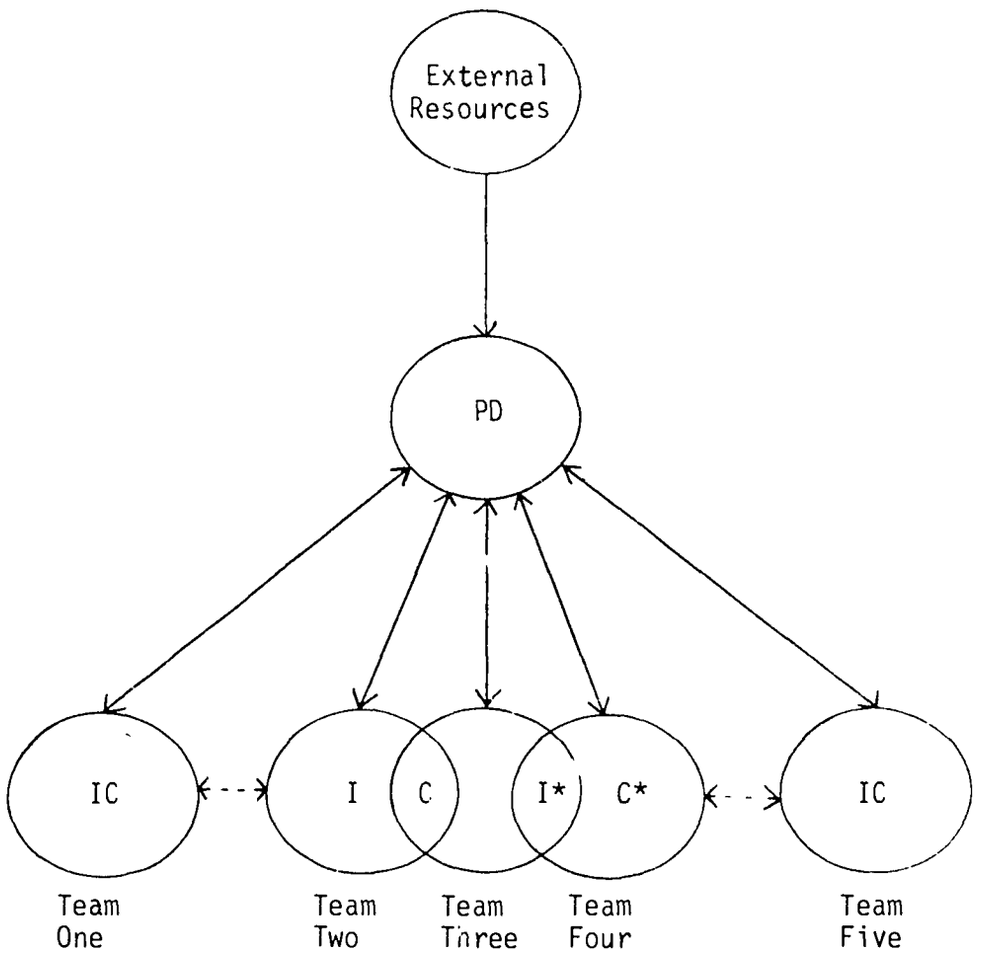


Figure 1.--Management/communications network

Instructional Design Procedure

A flowchart is presented in Figure 2 to summarize the instructional design procedure used to develop the modules. Although all of the finished modules were similar in form, slightly different problems were encountered during the development of each module. To reflect the fact that the design procedure varied across modules, the flowchart is presented in a "branching" format where the branches indicate alternatives which were followed in response to specific problems.

The first step in the instructional design process consisted of an assessment of student needs. The assessment was basically an attempt to answer two questions: What information and/or competencies that are valuable for secondary school teachers to possess are not being taught within the current curriculum? What information and/or competencies currently being taught to secondary school teacher candidates can be taught more effectively and efficiently using computer-managed instruction? (For a complete discussion of the identification of needs, see Gage, 1970). These questions were discussed with the class instructors, with the faculty course coordinator and in meetings of the instructional design group. As a result of this needs assessment, five topics were selected: (1) the use of computers in education; (2) the use of behavior modification techniques for classroom management; (3) the relationships between cultural differences and achievement in the school environment; (4) the use of teacher-made and standardized tests and principles of measurement; and (5) the use of statistical tools to interpret test scores. A constraint placed on the design was that the modules require no more than an average of four hours apiece in student study time.

After the topics were selected, the final two graduate students acting as content matter experts were selected and students were divided into the five instructional design teams. This division was made on the basis of competencies and interests. An instructional designer was paired with a content matter specialist whenever possible, thus allowing for the most effective use of the available talent. Content matter specialists were not required to spend time learning instructional design skills, and instructional designers were not required to spend time learning content matter. Thus, the skills which each individual brought to the project were immediately usable without the need for extensive training.

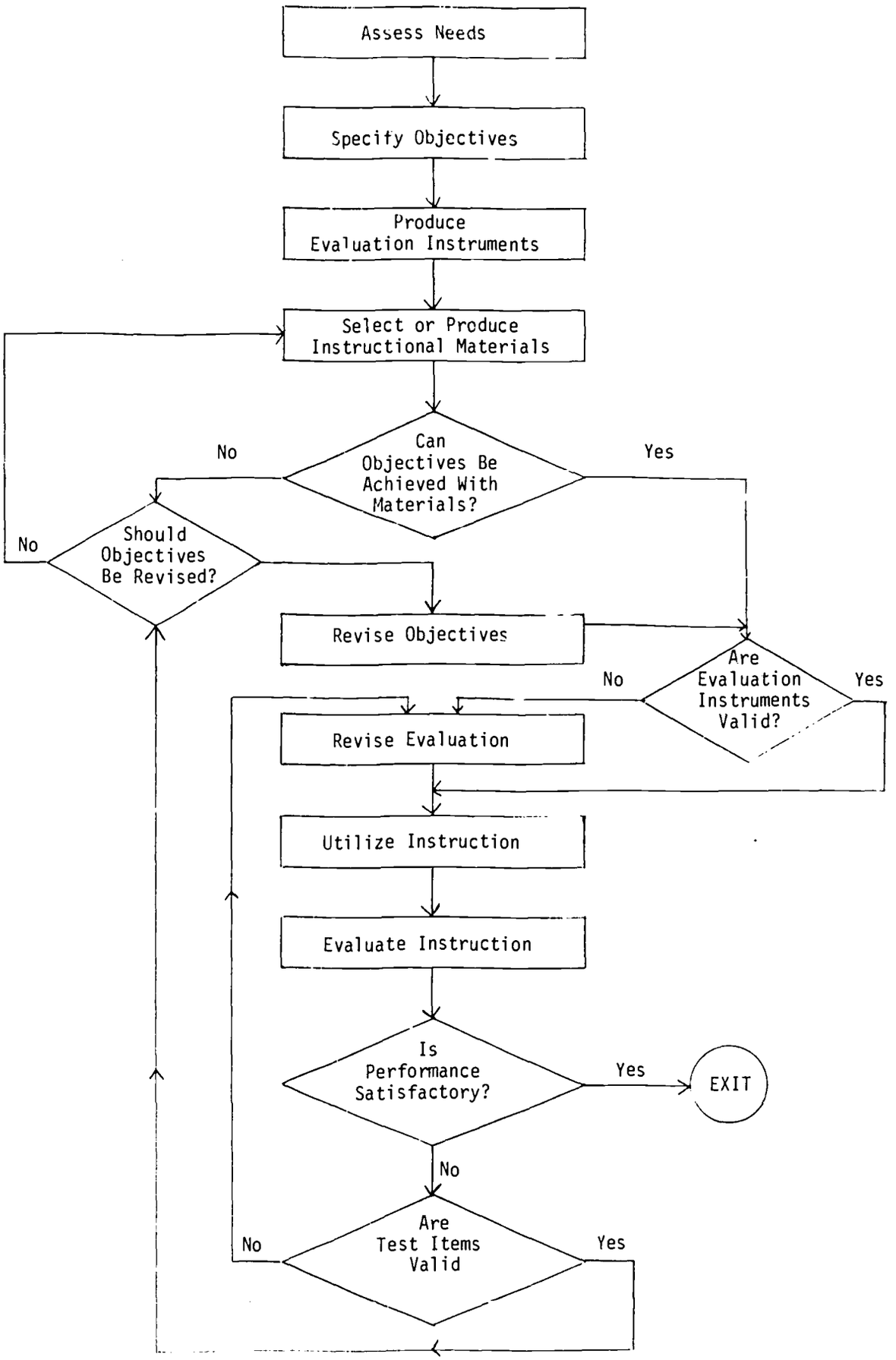


Figure 2.--Instructional design system flowchart.

Each team was given the responsibility for analyzing their specific topic and determining a set of behavioral objectives which could be achieved within a computer-managed instructional module. This specification of objectives relied heavily on the judgment of the content matter specialist. The instructional designer assisted the content specialist in stating the objectives in behavioral terms. The designer also classified each objective as being either a knowledge, comprehension, application, analysis, synthesis, or evaluation objective in order to be sure that certain categories were not being overemphasized. The criteria presented by Bloom (1956) were used for this classification.

Once the objectives had been defined, the teams again met as a group. Each team presented and defended the objectives which it had formulated. Both the value of each objective and completeness of each set of objectives were considered. The objectives for each module were revised until the group agreed that they were satisfactory.

During the next stage each team wrote three to six multiple choice test items to evaluate mastery for each objective in their module. The teams then met again as a group to review and criticize each other's work.

Following the specification of the objectives and evaluation test items, the next task was to secure the instructional materials. Generally, the first step was to search for suitable materials available at little or no cost. For example, the team which designed the measurement module selected some materials which were available at no cost from Educational Testing Service (1971). If available materials were inadequate, the team designed and produced a draft of the instructional material which they felt was necessary for the achievement of the stated objectives.

After materials had been secured or developed, the teams again met as a group to evaluate the materials selected. The critical question asked was whether the specified objectives could be achieved with these materials. If the answer was "no", a decision was made regarding whether to change materials or revise the objectives. Since these decisions had to be made on the basis of "expert" judgment, a group decision was considered preferable to decisions by individuals in order to insure a higher degree of accuracy (e.g., Smith, Stanley, & Shores, 1957).

Once the instructional materials had been selected, the validity of the evaluation items was reconsidered in order to avoid the possibility that the materials were teaching only the test rather than the desired behavior. For example, in order to test the achievement of an objective requiring the application of a principle, the student must be required to apply the principle in a novel situation (Bloom, Hastings, & Madaus, 1971). If a test item requires the application of the principle to a familiar situation, the item has no validity. Items judged to be invalid by the instructional designers were replaced. Since this judgment was considered to be rather straightforward, each team established the content validity of their own items. The items were then divided into three sets. The other two sets were used to produce two equivalent forms for evaluating student performance following instruction.

Pilot Evaluation and Revision

At this point, the modules were ready for an initial, small-scale trial. Each of the four modules was administered to one or two classes of students (n = 30 per class) registered in the Educational Psychology course during the 1972 summer session. The fifth module, Cultural Differences, was not completed in time for a summer trial. During this initial trial, the evaluation instruments were administered via paper and pencil tests and each student was presented with all items from both forms of the test.

Test results and student attitude data gathered by means of a questionnaire were used as a basis for revising objectives, instructional materials, and test items. The test scores obtained were somewhat higher than had been anticipated but this was considered to be at least partially an artifact of mode of test administration. The test items had been developed for computer administration, in which case the student would see only one item at a time and would not be able to skip an item and return to it later. This minimizes the possibility that information in one item can serve as a cue to the answer of another question. In contrast, students could skip items and return to them later in the paper and pencil form of the test. In addition, the course instructors maintained that summer students appeared to be somewhat different from students taking the course during the regular term, and, thus, revisions were made with caution.

In general, about 20 percent of the instructional materials and 30 percent of the test items were replaced or underwent some degree of revision. The format in which the modules were presented was modified slightly and an introduction to the set of modules was written. Factors which seemed as though they might become problematic were identified. For example, while certain materials were only moderately effective, the data did not clearly indicate that these materials should be replaced or revised. Therefore, such materials were retained for further evaluation.

Final Module Format

Following revision, the instructional modules and an introduction to the modules were published "in house" as Educational Psychology Modules, 1972. The format of each module was similar and consisted of an introduction to the module, specific reading assignments, each with a short overview of the reading assigned, behavioral objectives and sample test questions.

The introduction was designed to function as an advanced organizer. The assigned readings, of course, contained the material which the student is expected to learn. For the most part, the readings assigned were contained within the modules themselves. However, the majority of the reading materials for the Classroom Management and Statistics modules were contained in textbooks to be purchased by the student. A portion of the Tests and Measurement module was contained in materials supplied by the Educational Testing Service. The reading materials for the Computers in Education module were supplemented by a short demonstration of computer-assisted instruction. This demonstration also contained instruction on the operation of the terminal and served to familiarize the student with the facilities and procedures of the Computer-Assisted Instruction Laboratory.

The objectives specified, in behavioral terms, the competencies which the student was expected to exhibit following instruction. Sample test questions keyed to the objectives served to further define the objectives for the student and to familiarize him with the type of instrument by which he would be evaluated.

In order to provide a more complete description of the content of the modules, the introduction, behavioral objectives, and sample test items for each module are presented in Appendices A through E. The material is presented exactly as it appears in Educational Psychology Modules, 1972. The reader who is unfamiliar with modularized instruction is encouraged to examine these appendices before proceeding.

TEST PROGRAM AND DATA MANAGEMENT SYSTEM

Test Program Development

The final form of the evaluation instruments was not determined until after analysis of the results of the summer session trial. Two parallel test forms were defined for each module, each containing one to four questions per objective. The primary motivation for the construction of parallel forms was to provide a means by which each student could be tested twice over the same material, either on a pretest-posttest basis or on a posttest-retest basis. A secondary consideration was to increase test security. The availability of different forms allowed students to be randomly assigned to one of the two forms. As a further test security precaution, two different sequences of items were defined within each form.

The number of questions per objective was determined on the basis of the module authors' rating of the importance of each objective. While it would have been desirable to include at least four items per objective, total test length was constrained by the amount of time which students could be asked to devote to evaluation. With the exception of statistics, each module test form contained from 15 to 30 items. The statistics module contained a total of 60 items per form, but the complete module test was split into five subtests corresponding to the five units within the module. The subtests were administered to students one at a time and were treated as separate tests for programming purposes. All test questions were of a multiple choice format and, with only two exceptions, each contained four alternatives.

Master copies of the test questions and their correct alternatives were typed and kept on file. Each test question was formatted on a planning guide representing the character spaces available on the IBM 1500 system cathode-ray tube displays. Cards punched directly from these planning guides served as input to a preprocessor program run on the CDC 6600 computer system. This program produced Coursewriter II code for the IBM 1500 system. (The preprocessor program is described in Mitchell & Conner, 1971.) The code

produced by the preprocessor was then integrated into the control program written by the 1500 system programmers and assembled. Due to the program's relatively simple decision logic and the use of the preprocessor, only a limited amount of on-line debugging time was required. The test program control logic and the characteristics of the data management system are described in the following section of this report.

In their programmed form, test items for a particular module were presented one at a time on a cathode-ray tube display. Students responded by typing a 1, 2, 3, or 4 corresponding to the response alternative selected. A response other than one of these four characters elicited a message to the student requesting him to type only a 1, 2, 3, or 4.

Students were given diagnostic feedback on their performance immediately following their completion of each test. This consisted of the student's total score on the test (as a percentage), the criterion percentage set by his instructor, and his percentage score on each objective. A typewritten copy of the same feedback was also available if the student requested it. Two types of reports summarizing student performance were available for the class instructors. These reports were also prepared on the typewriter terminals and were available to the instructors within a few minutes after being requested. The characteristics of these reports are described in more detail below.

Data Management System

The goals directing the design of the CMI data management system were to provide flexible, on-line testing for a large number of students and to facilitate on-line report generation. Since the conventional 1500 system response data storage and retrieval process involves the relatively cumbersome use of magnetic tape, the requirements of CMI dictated the use of disk for bookkeeping and student response data storage.

In the current version of the disk centered system, data for up to 640 students are stored on a disk pack and accessed for on-line summaries. The file structure is designed to accommodate up to three class sections for

each of up to 20 instructors. Each class section may contain up to 60 students, subject to the overall constraint of 640 students. The information contained on the CMI data management recording pack is of four types: bookkeeping information, course configuration information, section configuration, and student data.

Bookkeeping information consists of pointers to available files, a current count of the instructors and sections registered on the data pack, and a current count of the number of students registered over all class sections. This information is accessed during initialization of the data pack, i.e., instructor, section and student registration, prior to use of the CMI tests by students. The structure in which this information is stored is represented in Figure 1 of Appendix F.

Course configuration information specifically defines the content of each CMI module to be administered. Two sectors on the disk pack (see Figure 2, Appendix F) are allocated for this information. The first word of the first sector indicates the total number of modules (up to 25) available. Sequentially thereafter are a set of up to 25 26-word records which specify the number of objectives and the number of test items per objective for each module. A given module may consist of up to 20 objectives and each individual objective may have up to ten test items. Course configuration information is utilized in determining performance (percent correct) for a given objective within a module and in determining overall performance for the module. This information is also used in producing the summary printouts for instructors and in providing performance feedback to students. The course configuration is specified on-line by a proctor through the use of a program written in Coursewriter II specifically for setting up of the data pack and for on-line retrieval and summary of student data.

Section configuration information specified the module tests which are to be administered to students in a given class section, the sequence in which these tests are to be administered, and the performance criterion assigned for mastery for each module. Thus, each section can have a module assignment and sequence which is unique to that section together with a

unique set of performance criteria. This allows the use of a single instructional program containing all modules to be used for a number of class sections with differing requirements. The instructional program's control routine selects only the module tests specified (and selects them in the appropriate sequence) for administration to a student in a given section. Each class section is identified by a two character identifier--the first two characters of the student number. The structure in which this information is stored is outlined in Figure 3 of Appendix F. Section configuration information is used to determine whether or not a student has achieved the specified criterion both at the objective and module levels. This information is also used in producing summary printouts for the instructors and in giving performance feedback to students.

Student data stored on the data pack consist of a set of eight-word records in the format shown in Figure 4 of Appendix F. Within a record, raw scores are maintained for each objective in a module together with latency data for the module test, and date (month and day) of test administration. During the formative evaluation, data from the curiosity and anxiety scales administered with each module were also maintained with each record. An otherwise unused portion of the recording pack was used to store student names which could be accessed positionally. One eight-word record is written to disk upon completion of a test or retest. Records for one initial test and up to two retests may be stored for each module. The series of records written to disk for a given student are all contained within a two-sector block reserved for that student. Student data records are utilized in producing summary reports for the instructors.

During the formative evaluation, allocation of each student performance record block was handled during section registration and configuration. The maximum number of students anticipated for a given section had to be entered by the proctor at section registration time. From this information and certain bookkeeping information, an N-word table (where N is the number of students) of pointers to recording blocks was set up for that class section.

Thus, when a student in a given section initially signed onto the system, it was only necessary to determine his section (from the first two characters of his student number) and to find the pointer to the allocated recording block by indexing the table of pointers using the binary value of the last two characters (digits 01-60) of his student number.

The requirement that the number of students in each section be entered prior to the time the students signed onto the system was found to be unsatisfactory during the formative evaluation due to the number of students adding and dropping the course after registration. The scheme has since been modified to dynamically allocate blocks at the time of initial student sign on.

As was mentioned above, one of the primary purposes of the CMI data management system was to facilitate the generation of timely and meaningful reports to instructors. This is done by use of a "course" coded in Coursewriter II which performs report generation as one of its functions. Reports to instructors may be generated at two levels: a course summary report of information at the module level and a module summary report of information at the objective level.

A course summary report is specific to a given class section and a given instructor. The report is headed by the instructor's name and the section number, followed by a listing of all modules to be taken by students in that section in the sequence specified in the section configuration record. The criterion score for each module is also given. The body of the report is a student-by-module score (percentage) for all students in the section over all modules taken up to the time of the report. Students are identified by both name and student number. The percentage scores shown for each module are scores after retest if a retest was administered. An "R" is appended to the test score if a retest has been administered.

A module summary report is specific to instructor and section as above but is also specific to a particular module. The report is headed by instructor name, section number, module number and module criterion score,

followed by a listing of the total number of test items for each objective and the number of correct answers required to meet criterion for each objective. The body of the report is a student-by-objective score (number correct) for all students over all objectives in the module. Students are identified by both name and student number. Raw scores are shown for both initial posttest and retest, if a retest was administered. Retest scores are shown only for those objectives taken on the retest (those on which the student did not reach criterion). Test administration dates (day, month) are shown for both test and retest.

All disk reads/writes, both in the instructional program and the data pack program "set-up" and data summary operations are performed through an assembly language function, DISKIØ, originally written by Thomas McMurchie at Florida State University. It was modified by The University of Texas Computer-Assisted Instruction Laboratory staff to provide the capability of treating counters and buffers as arrays for transfer and thus reducing the number of reads/writes required to transfer information to and from disk.

FORMATIVE EVALUATION

Student Sample

The first broad scale use of the modules occurred during the Fall, 1972, semester. It was initially assumed that approximately 200 students would be run, this number being considered adequate to provide a data base for evaluation and revision. The course coordinator and instructors, however, requested that the modules be made available to all students registered in the course. Consequently, the modules were used by approximately 560 secondary education students enrolled in the Educational Psychology 332S course. Since these students are secondary education trainees, entering the formal education program during their junior year, their lower-division courses represent a variety of different disciplines.

The students were enrolled in 15 different sections of 332S taught by 14 instructors. Five of the 15 sections were already involved in an experimental program entitled "Proctor-Assisted Instruction." These five sections differed from the ten regular sections in terms of classroom instructional techniques and class schedule. As a result of differing requirements of these two groups, they were treated separately for purposes of scheduling and evaluation.

Procedures

The 330 students in the 10 conventional classes received the instruction during September and early October, 1972. These students studied only four of the five modules since their instructors did not wish to include the Statistics module in the curriculum. The evaluation of the Computers in Education, Classroom Management, Tests and Measurement, and Cultural Differences modules reported here was based on these students.

The 200 students and 30 proctors in the five proctor-assisted instruction classes studied all five modules. These students received the instruction during October, early November, and the first week of December, 1972. The only data from these students to be discussed in this report concern the evaluation of the Statistics module and pretest data for all modules. Data from the proctors have not been used for evaluation purposes and will not be reported.

Pre-Instructional Evaluation. While it was desirable that the test item pool be evaluated through administration to students prior to their receiving instruction, a conventional pretest-instruction-posttest procedure was not considered to be feasible. Only two test forms were available for each module and it was considered preferable to reserve both forms for post-instructional testing and retesting. The additional demands which conventional pretesting would place on both student and computer terminal time were also considered to be undesirable. The decision was made, therefore, to administer a paper and pencil pretest battery to a sample of students drawn from the 332S classes. The desire to maintain a degree of test security raised additional problems for pretest administration. While the simplest procedure would have been to administer complete tests from the different modules to individual students, this would have given each student tested complete exposure to one form of one module test. The approach adopted involved a matrix sampling procedure in which each student was administered a test consisting of items drawn from all tests.

The total pool of 298 items was divided into six subtests of 49 or 50 items each. Each subtest contained items drawn from both forms of all module tests and all subtests contained approximately equivalent numbers of items from each of the module tests. An individual test item appeared on only a single subtest and no subtest contained parallel items from the alternative test forms for a single module.

The pretest was conducted with 100 students drawn from three of the five proctor-assisted instruction classes during the month of September, prior to their introduction to the modules. Testing was conducted in groups and was completed within a five-day period. Each student tested was randomly assigned to one of the six subtests. Fifteen to seventeen students completed each subtest, thus providing fifteen to seventeen responses to each of the 298 items in the total test item pool.

Instruction and Posttest. Details of the procedures to be employed during the formative evaluation were determined in cooperation with the class instructors. Each student was required to attain a certain level of mastery on each module rather than on each objective within a module. Specifically, students were required to achieve a score of 75% on each module. If a student failed to

reach this criterion on his initial posttest, he was administered a retest based on the alternative test form. For all modules other than Statistics, the retest items covered only those objectives on which the student scores less than 75% on his initial posttest. His final module score, following the retest, was a composite of his scores on objectives passed on the initial test and his retest scores on objectives failed on the initial posttest and retested. Since the Statistics module subtests were quite short, the decision was made to retest over all objectives in a subtest following a student's failure on that subtest. If a student failed to achieve criterion following a retest, he was referred to his instructor for an oral examination.

Students in the ten conventional classes were introduced to the modules during the first regularly-scheduled class meeting. The manual containing the instructional modules (Educational Psychology Modules, 1972) and the materials from Educational Testing Service for the Tests and Measurement module were distributed (at no charge) and the students were given a brief lecture on Computer-Managed Instruction in general and the procedures to be followed in this particular instance. An outline of these procedures was also provided in the manual's introduction to the modules. Each instructor informed his class of the criterion established for mastery, of the dates by which each module test was to be completed, and that the content of the modules would not be discussed in class prior to the completion date.

Approximately one week was allowed for the completion of each module. An extra three days was provided for the completion of the first module, Computers in Education, to allow the students more time to become familiar with the procedures. The order in which subsequent modules were assigned was as follows: Classroom Management, Tests and Management, and Cultural Differences. Essentially the same procedures were followed for the five proctor-assisted instruction classes, with the exceptions that they were introduced to the modules one month later, and modules were assigned in a slightly different order: Computers in Education, Statistics, Tests and Measurement, Classroom Management, and Cultural Differences.

Students were assigned to one of the two test forms on the basis of the student numbers by which they were identified by the computer program.

Students with numbers ending in an odd digit were assigned to test form A for their initial posttest on each module while students with numbers ending in an even digit were assigned to form B for their initial posttest. Students were assigned to one of the two different item sequences within each test form on a random basis.

The students' first contact with the computer terminals was for purposes of a demonstration. This demonstration, which was treated as a component of the Computers in Education module, consisted of a short computer-assisted instruction tutorial on the use of the computer terminals and a brief demonstration of a segment of a computer-assisted instruction program on English punctuation and grammar. To facilitate scheduling, and since no data were being collected, students were encouraged to sign up for demonstration appointments two at a time.

The eight cathode ray tube (CRT) and typewriter keyboard terminals in the Laboratory terminal room were available seventy hours per week: noon until 10:00 p.m. on Mondays, 9:00 a.m. until 10:00 p.m. on Tuesday through Friday, and 9:00 a.m. until 5:00 p.m. on Saturday. Terminal time for the initial demonstration was scheduled in one-hour blocks. Terminal time for tests and retests was scheduled in one-half-hour blocks.

Having received his instructional materials and instruction on the use of the terminal, a student began his study of the first module. When he thought that he could achieve the stated objectives for the module, he telephoned the computer terminal room to schedule a 30-minute block of terminal time. During the initial few tests, a student was signed onto the terminal by a proctor. During later tests, students signed themselves on. The first message displayed following sign-on was the students' name and a message which asked him to call the proctor if it was not his name which was displayed. This was to assure that students signing themselves on had entered their correct student number.

After signing onto the terminal but prior to beginning a test, the student was administered an on-line version of a nine-item curiosity scale (Leherissey, 1972) concerning his interest in the reading materials. The module test items were then presented one at a time and the student responded by typing a 1, a 2, a 3, or a 4. Immediately following his completion of

the test, the student was administered a five-item state anxiety scale (Spielberger, Gorsuch, & Lushene, 1970) asking him about his emotional state during the test. The student was then given feedback on the CRT display regarding his performance. The feedback display included the criterion percentage established by his instructor, his own percentage score over the complete module, and his percentage scores on the individual objectives. Objectives for which the student achieved or surpassed the criterion percentage were starred. The student was then asked whether or not he wanted a typewriter printout of the diagnostic report. If the student achieved criterion, he was finished with that module. If he failed, he was instructed to restudy the materials and to make an appointment for a retest. On a random basis, one quarter of the students were asked to complete a thirteen-item paper and pencil attitude questionnaire following each of the four modules.

In general the same procedures were followed for retests with the exception that the curiosity scale was not administered, the student was tested on only those objectives failed on the initial test, and no attitude scales were administered.

Evaluation Results

Pretest Results. Data resulting from the pre-instructional evaluation are shown in Table 1. Recall that since no one student was administered more than one-sixth of the items from any one module test, the total test scores shown represent composites of the performance of all students in the sample. Scores are presented in terms of percentages in order to facilitate comparisons among tests and across objectives. The five subtests from the Statistics module have been presented separately.

In general, the obtained scores were higher than had been anticipated, much higher, of course, than would have been expected strictly on the basis of chance (25%). To some extent, the scores were probably inflated (relative to subsequent posttest scores) due to the use of a paper and pencil test as opposed to testing at the computer terminal. Even when this potential error term is taken into consideration, it must be concluded that the tests contained some poorly constructed items to which the answers were obvious or that some

Table 1

PRE-INSTRUCTIONAL EVALUATION OF FIVE CMI MODULE TESTS

(SCORES SHOWN ARE COMPOSITE PERCENTAGES. N = 17)

Module	Percentage Correct on Total Test	Percentage Correct on Each Objective*											
		1	2	3	4	5	6	7	8	9	10	11	12
Computers in Education	55	55	66	50	50	52	56	48					
Classroom Management	53	46	45	42	68	53	62						
Tests and Measurement	55	61	53	70	54	60	17	57	57	46	47	62	79
Cultural Differences	48	57	54	41	63	20	39	51	40	60	60	29	64
Statistics, Unit I	51	78	40	55									
Statistics, Unit II	33	38	29										
Statistics, Unit III	50	62	43	40	42								
Statistics, Unit IV	49	77	30	37	59	48							
Statistics, Unit V	37	32	53										

*The number of objectives varied among modules.

students already possessed skills to be taught in the instructional modules. Undoubtedly both of these factors were present. Despite the fact that the obtained scores were somewhat higher than had been anticipated, the results did indicate that, in general, the performance of this sample of students was well below the desired level of competency.

An item analysis of the data was conducted which provided the probability of a correct response for each item, the item's point bi-serial correlation with total test score and the distribution of responses across the four alternatives. Separate response distributions were also provided for high and low scoring students based on a median split of total subtest scores. While these detailed data will not be reported, they were used to provide one basis for revision of the module tests.

Posttest Results. The percentages of students in the ten conventional classes who achieved the criterion of 75% correct on each of four modules are shown in Table 2. As may be noted, there was some attrition (19%) from the first to the last module. For the most part, this was due to students' dropping the course during the first month of the semester, although there were a few students who did not complete the last posttest prior to the date on which data collection for this report was terminated.

The percentage of students achieving criterion on the initial posttests was lower than would eventually be desired, particularly for the fourth module, Cultural Differences. In general, however, it was considered to be satisfactory in view of the modules' stage of development. Student performance on retests was at least as good as had been anticipated and as a result the percentage of students eventually achieving criterion, on either the initial test or the retest, was quite satisfactory for three of the four modules.

The Cultural Differences module, while definitely in need of revision, probably suffered to some extent from its placement at the end of the sequence of modules. Since the final deadline for completing the modules was approaching, many students probably allocated less time to studying this module than they had to the others. Unfortunately, the Cultural Differences module contained the largest amount of reading materials of any of the modules. The

Table 2

PERFORMANCE OF TEN CONVENTIONAL CLASSES ON FOUR EDUCATIONAL PSYCHOLOGY MODULES

(CRITERION = 75% CORRECT)

Module	Performance on the Initial Posttest		Performance on the Retest		Overall Performance	
	Number of Students	Percentage Achieving Criterion	Number of Students Taking Retest	Percentage Achieving Criterion	Number of Students Achieving Criterion	Percentage of Students Taking the Posttest Who Achieved Criterion
Computers in Education	332	69	100	76	306	92
Classroom Management	315	75	76	76	297	94
Tests and Measurement	302	67	92	82	280	92
Cultural Differences	270	53	79	83	211	78

effects of the deadline are also indicated by the fact that while 125 students failed to achieve criterion on the initial posttest on Cultural Differences, only 79 of these students scheduled a retest prior to the final data collection date. This discrepancy partially accounts for the relatively low percentage of students finally achieving criterion, 78%. It should be noted that Cultural Differences was the one module which had not been pilot tested and revised during the summer session.

The percentages of students in the five proctor-assisted instruction classes who achieved the criterion of 75% correct on each of the five subtests in the Statistics module are shown in Table 3.

The percentages of students achieving criterion on the initial posttests are comparable to the percentages for the other four modules shown in Table 2. The percentages of students achieving criterion on retest are somewhat lower than the comparable percentages shown in Table 2. This discrepancy may be due to the nature of the subject matter or to the fact that students taking Statistics were retested over all objectives in a unit rather than over only the particular objectives failed on the initial posttest. This possibility will be explored in subsequent research. The final percentages of students eventually achieving criterion, on either the initial posttest or the retest, were considered satisfactory in view of the stage of the module's development.

The mean percentage of items answered correctly on each of the initial posttest is shown in Table 4. Data are shown for both individual objectives, for the total test for each module, and for each unit subtest for the Statistics module. The magnitude of pre- to post-instructional gain may be inferred by comparing these data to the data presented in Table 1, but it should be remembered that the data in the two tables do not represent the same set of students.

Equality of Forms and Sequence. It will be recalled that two alternative test forms were developed for each module. Alternative forms were assigned to students on a pseudo-random basis; that is, on the basis of the last digit of their student number. With the exception of the Statistics module, items within each test form were also presented in two different orders.

Table 3

PERFORMANCE OF FIVE PROCTOR-ASSISTED INSTRUCTION CLASSES ON THE FIVE UNIT TEST IN THE STATISTICS MODULE (CRITERION = 75% CORRECT)

Unit Test Number	Performance on Initial Posttest		Performance on Retest		Overall Performance	
	Number of Students	Percentage Achieving Criterion	Number of Students	Percentage Achieving Criterion	Number of Students Achieving Criterion	Percentage of Students Taking Post-test Who Achieved Criterion
1	198	58	82	60	165	83
2	197	64	69	65	172	87
3	196	89	20	75	190	96
4	196	80	38	73	185	94
5	194	70	53	67	173	89

Table 4

MEAN PERCENTAGES CORRECT ON INITIAL POSTTESTS

Module	N	Test Total	Objectives 1	2	3	4	5	6	7	8	9	10	11	12
Computers in Education	332	81	86	85	95	73	75	95	72					
Classroom Management	315	79	79	75	89	75	75	80						
Tests and Measurement	302	78	92	83	80	79	75	77	84	74	76	80	75	90
Cultural Differences	270	73	73	83	77	74	69	60	79	67	66	84	65	82
Statistics, Unit I	198	77	84	74	77									
Statistics, Unit II	197	77	73	81										
Statistics, Unit III	196	81	71	82	57	83								
Statistics, Unit IV	196	83	96	82	57	83	84							
Statistics, Unit V	194	77	78	77										

Only a single sequence of items was developed for the Statistics tests due to disk storage limitations. While the alternative forms and sequences were designed to be of equal difficulty, an empirical check on this assumed equality was obviously necessary. For each of the first four modules, the total scores obtained by students assigned to each of the two item sequences and each of the two forms were subjected to a two-by-two analysis of variance. Scores obtained by students assigned to each of the two forms of the Statistics post-test were subjected to a one-way analysis of variance.

No significant differences were found between item sequences for any module. No differences were found between test forms for the Computers in Education or Classroom Management modules. Within the Tests and Measurement module, form B was found to have a lower mean score (22.4) than form A (24.5). This difference was significant ($p < .001$). Similarly, form B was found to have a significantly lower mean score (17.0) than form A (18.2) within the Cultural Differences module ($p < .01$). Differences in difficulty were found between forms in two of the five-unit subtests in the Statistics module: Units I and II. In Unit I, form A was found to have a significantly lower mean score than form B (4.1 as opposed to 5.1) ($p < .01$). Form A was also found to have a significantly lower mean score than form B in the Unit II subtest (4.2 as opposed to 5.0) ($p < .001$). These differences were taken into consideration during revision of the modules in that test items for which performance data were available were distributed across forms so as to balance difficulty of individual objectives.

Scheduling Considerations. A variety of time factors were also considered in the evaluation of the total CMI system. Specifically, the time required for students to complete their study of the modules, the terminal time required for students to complete the module posttests, and the pattern of scheduling for the posttests were determined. These data were based on only the 330 students in the ten conventional sections of the course and on the four modules studied by these students.

It will be recalled that one of the constraints placed on the design of the instructional modules was that, on the average, the time required for a student to complete his study of a module not exceed four hours. Since the

instructional system did not allow a means of directly monitoring the amount of time students spent on the different modules, one item on the module questionnaire asked the student to estimate the amount of time which he devoted to studying the module. The specific question was stated as follows:

> How much time did you need to read all of the instructional materials for this module? Do not include the time for the CAI demonstration in the first module.
>
> a. 1 hour or less.
> b. 1 to 2 hours.
> c. 2 to 3 hours.
> d. 3 to 4 hours.
> e. 4 hours or more.

The questionnaires were administered to approximately one-quarter of the students following their completion of each module posttest. A summary of their reported study times is shown in Table 5.

Table 5

DISTRIBUTION OF REPORTED STUDY TIMES

TO COMPLETE EACH MODULE

Module	N	Percentage of Students Reporting Each Amount of Time			
		< 2 hours	2-3 hours	3-4 hours	> 4 hours
Computers in Education	63	30%	35%	25%	10%
Classroom Management	70	13	26	29	32
Tests and Measurement	74	9	26	28	37
Cultural Differences	56	2	11	25	62

The Computers in Education module was obviously not too long: 90% of the students completed the module in four hours or less. The Classroom Management and Tests and Measurement modules were marginal in that approximately one-third of the students reported that they required more than four hours of study. The Cultural Differences module was definitely too long, with only one-third of the students reporting that they completed their study of the module in four hours or less.

A second consideration involved the amount of terminal time required for students to complete the various module posttests. This information was available from the computer system records. Since the lengths of the retests were variable and never exceeded the lengths of the corresponding initial posttests, only initial posttest times were considered. The distributions of student terminal times required to complete the four modules are summarized in Table 6. The terminal times reported include the time required for the student to complete the nine-item curiosity scale and the five-item anxiety scale as well as the test itself. No systematic attempt was made to isolate the time required for students to complete these scales but it is estimated that they required a total of four to seven minutes of the students' time.

Table 6

MODULE POSTTEST TERMINAL TIMES: PERCENTAGE OF STUDENTS COMPLETING TEST WITHIN EACH TIME PERIOD

Module	No. of Questions	N	Time in Minutes				
			< 15	15-20	20-25	25-30	> 30
Computers in Education	15	332	89%	8%	2%	1%	0%
Classroom Management	20	315	48	29	15	4	4
Tests and Measurement	30	302	13	29	27	20	11
Cultural Differences	24	270	8	24	30	19	19

Student terminal time was, of course, related to some extent to the number of questions on the test. It will be recalled that terminal time for the tests and retests was scheduled in 30-minute blocks. The shortest two tests, Computers in Education and Classroom Management, were completed by almost all students well within the scheduled block of time. The small percentage of students who required more than 30 minutes to complete these tests was easily absorbed in the slack in the scheduling procedures. Larger but still tolerable percentages of students required more than 30 minutes to complete the Tests and Measurement and Cultural Differences posttests. Both of these tests were shortened somewhat during revision.

It is interesting to note that there was a tendency for students to require more time to complete the Cultural Differences posttest than to complete the Tests and Measurement posttest, despite the fact that the latter contained 25% more test items. It will be recalled that students registered the poorest posttest performance on the Cultural Differences module. The higher relative difficulty of this test was reflected in the students' longer response latencies to the test items and thus in their increased terminal time.

In view of the fact that a large proportion of a CMI system's operational costs are directly attributable to student terminal time, it would be desirable to reduce the amount of time required for posttests. To maintain a workable schedule, the next shorter time period which would be considered would be 20 as opposed to 30 minutes. Such a schedule would allow three testing sessions per hour rather than the current two, a one-third reduction in the required terminal time. Assuming that the curiosity and anxiety scales were deleted from the testing session or at least substantially shortened, it would appear that the maximum test length allowable would be approximately 20 items.

As was previously mentioned, students were given deadlines by which they were to have completed the posttest (and retest, if necessary) for each module. Initially, nine days were allowed for the Computers in Education module and six days were allowed for the completion of each of the three successive modules. Prior to the expiration of the second deadline, both the

instructors and students requested that it be extended. Consequently, all deadlines were extended slightly. Seven days were allowed for the completion of the second and third modules. The final deadline was constrained by the fact that the second group of students had begun their testing sequence. Consequently, this completion date was allowed to slip only one day, allowing just four days for the completion of the Cultural Differences module. It should be remembered that these dates were final deadlines. Students were free to take the posttest for a particular module prior to the completion date for the previous module.

The cumulative percentages of students who had taken each module posttest (and retest, if necessary) by a particular date are shown in Table 7. The figures shown in this table are percentages of the number of students who were administered each test and retest prior to the final data collection date. As may be seen from these data, there was a tendency on the part of the students to procrastinate. In general, a large number of students waited until the last or next to last day before the deadline to take the posttest. On the average, however, only about 10% of the students failed to schedule posttests prior to the assigned completion dates. Retest scheduling presented a more serious problem. Since a student could not know whether a retest would be required until he had taken his initial posttest, and since most posttests were scheduled relatively late in the assigned period, only about 60% of the retests were scheduled prior to the deadlines.

The most noticeable result of students' tendency to delay taking the posttests was the lack of regularity in terminal scheduling. The number of test sessions scheduled per day (both initial posttests and retests) over a one month period is shown in Figure 3. As is obvious from inspection of this figure, terminal usage was definitely not evenly distributed over time. Although terminal use never reached the maximum daily capacity of the system (208 test sessions), the four days on which the number of student tests exceeded 100 did raise problems of finding suitable appointment times for all students. Figure 3 represents a total of 1566 tests and retests. A steady level of about 75 tests per day would have permitted all of the students represented to have completed their testing by the final deadline without interfering with other activities using the computer terminals.

Table 7

CUMULATIVE PERCENTAGES OF STUDENTS COMPLETING TESTS AND RETESTS BY DATE

(VERTICAL DIVISIONS REPRESENT STATED MODULE COMPLETION DEADLINES)

Module		September											October																
		11	12	13	14	15	16	18	19	20	21	22	23	25	26	27	28	29	30	2	3	4	5	6	7	9	10	11	12	13
Computers in Education	Test	8	20	44	71	89	92	95	95	96	97	98	98	98	98	98	99	100												
	Retest	2	6	16	34	58	73	88	93	93	94	94	95	96	97	97	97	99	99	99	100									
Classroom Management	Test		0	0	1	3	4	11	24	45	76	89	92	94	95	96	96	97	98	98	99	99	100							
	Retest								3	17	32	61	70	83	88	93	95	95	96	96	97	97	99	100						
Cultural Differences	Test			0	0	0	0	2	3	5	5	8	11	16	29	46	65	84	89	94	96	99	99	99	100					
	Retest							1	2	2	2	2	2	7	10	23	35	54	74	88	95	97	97	98	98	98	98	98	98	100
Tests and Measurement	Test												0	1	4	5	6	9	12	21	50	91	91	93	94	94	96	98	99	100
	Retest															1	3	6	6	10	24	61	63	65	67	77	80	87	96	100

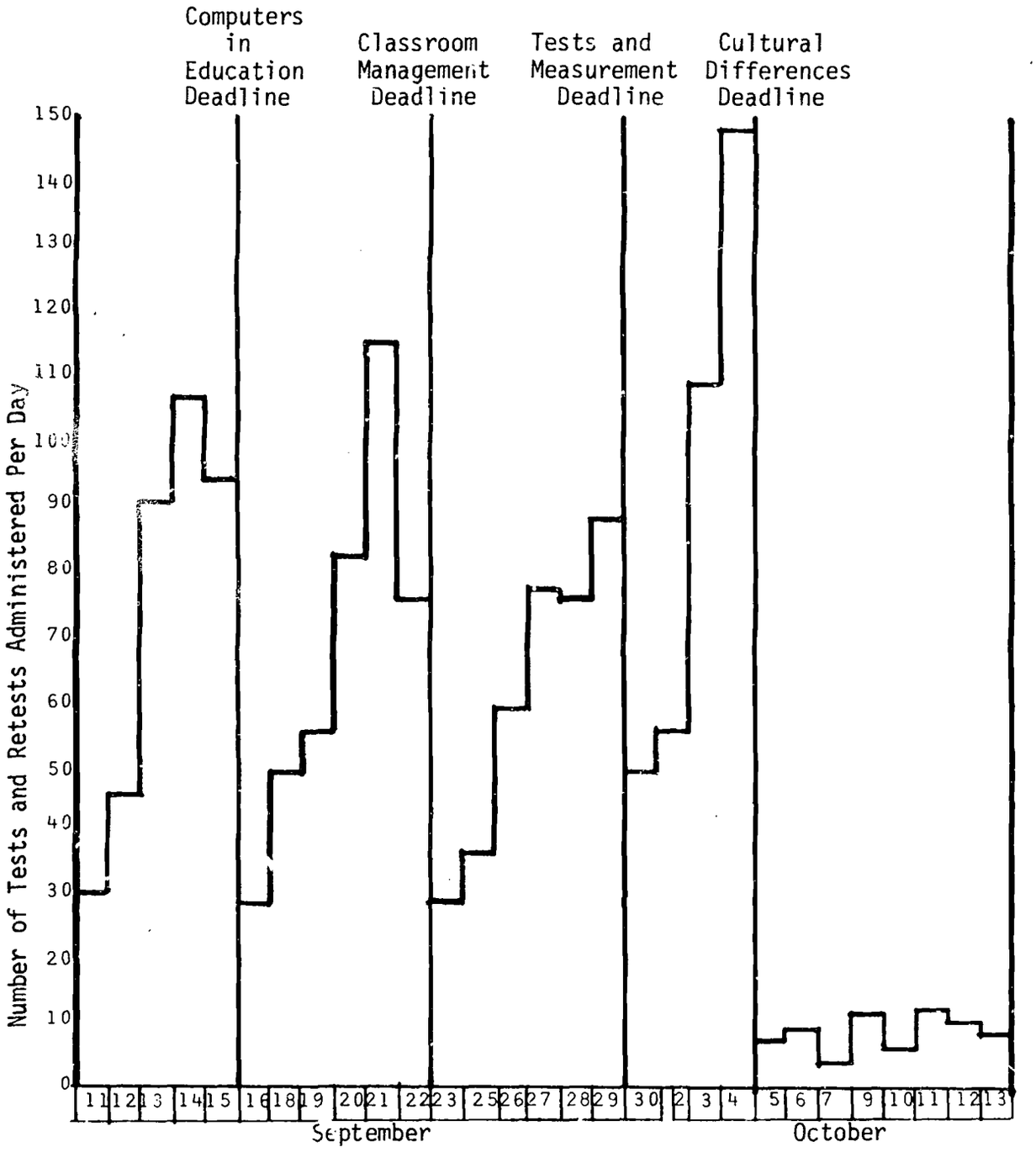


Figure 3.--Number of Tests and Retests Administered Per Day

The assignment of completion dates for individual modules would appear to be necessary for the context in which these modules were used. Given the pressures of other class assignments, it is not surprising that the students tended to procrastinate. In the future, however, the assigned deadlines for different classes will be staggered in an attempt to alleviate the very uneven terminal demands experienced.

REPORT OF ESTIMATED COSTS

Developmental Costs

The estimated direct costs of developing the data management system and the five computer-managed instructional modules are reported in Table 8. Indirect costs, such as fringe benefits and institutional overhead, have not been included since these would vary from one institution to another. Not all of the costs reported in Table 8 were incurred by the funding agency, some being absorbed by the institution and the individuals involved. These figures do not include the cost of revision following the formative evaluation.

Since personnel time constituted the major expense of instructional development (approximately 83%), a special procedure was employed in an attempt to insure the accurate recording of the expenditure of time. Each individual was given a form and asked to maintain a weekly record of the amount of time spent on the project. Several of the project personnel failed to maintain accurate records and the time expended by these individuals was estimated. Whenever possible, the accurately recorded time expenditures for the performance of similar tasks was used as a basis for these estimates.

Operational Costs

The operational costs associated with the use of four modules (excluding Statistics) for the 330 students in the ten conventional classes are reported in Table 9. Since procedures had been established for recording the costs of supplies and computer time prior to the implementation of this project, no special procedures were required to insure the accurate recording of these operational costs. Thus, the information in Table 9 was obtained through a simple analysis of the records normally maintained by the Computer-Assisted Instruction Laboratory.

Table 8

ESTIMATED DEVELOPMENT COSTS FOR THE DATA MANAGEMENT SYSTEM

AND THE FIVE INSTRUCTIONAL MODULES

	Estimated Time in Hours	Total Hours	Estimated Cost
Data Management Systems Development			
Personnel			
Management	35		
Systems Analyst	121	156	805.00*
Computer (IBM 1500 System)			
Author time @ $12.18/hr	25.9		
Utility time @ $48.73/hr	2.0	27.9	412.00
			1,217.00
Instructional System Development			
Personnel			
Management	415		
Authors and Instructional Designers	1246		
Systems Analyst	311		
Computer Programmers	415		
Keypunch Operators	35		
Secretarial/Clerical	380		
Instructors	87	2889	9,602.00*
Computer			
CDC 6600 System @ $260/hr	0.2		
IBM 1500 System:			
Author time @ $12.18/hr	44.3		
Utility time @ $48.73/hr	19.2	63.7	1,527.00
Copyright Fees			210.00
Supplies and Materials			100.00
			11,439.00

*Estimated salaries excluding fringe benefits

Table 9

CMI OPERATIONAL COSTS

(First 300 students using four modules)

	Time in Hours	Total Hours	Cost
Personnel			
Management	35		
Secretarial/Clerical	69		
Proctors	363		
Operators	190	657	$1,740.00*
Computer (IBM 1500 System)			
Student time @ $6.09/hr	822.0		
Utility time @ $48.73/hr	16.5	838.5	5,810.00
Instructional Materials			377.00
			$7,927.00

*Salaries excluding fringe benefits.

REVISION OF THE MODULES AND TESTS

Following collection of the formative evaluation data, the five modules all underwent substantial revision. This revision was based on three sources of data: (1) student performance from the pre-instructional evaluation; (2) student performance on the initial posttests; and (3) student responses to the attitude questionnaires administered following each module. In contrast to norm-references testing situations, no reliable procedures have been generally accepted for evaluating criterion-referenced tests or for diagnosing the problems underlying poor performance on these tests. Consequently the procedures followed in the revision process were more heuristic than algorithmic.

As a first step, the data from the pre-instructional evaluation were examined. If the mean probability of a correct response was found to be particularly high (greater than .60), that item was examined closely for construction deficiencies. This examination was guided by the pattern of responses to the distractor alternatives. Specifically, the following four questions were asked regarding the item: (1) Are the distractors plausible? (2) Are the distractors grammatically consistent with the stem? (3) Are there verbal associations between the stem and the alternatives? (4) Are the relative lengths of the alternatives equivalent? If no serious deficiencies were noted in any of the items assigned to a particular objective and if the mean probability of correct responses of items assigned to that objective was relatively high, the question was raised as to whether that objective should be included in the instruction.

A much larger data base was, of course, available for the post-instructional evaluation. In addition to the performance data described in this report, all of the students' responses to individual test items were recorded on magnetic tape. A card containing responses made to each item on his initial posttest was punched for each student for each module. All cards for each test form within each module were then submitted to an item analysis

program (Jennings, 1970) run on the University's CDC 6600 computer. This program divided the students into pantiles based on their total test score and indicated the number of students within each pantile selecting each alternative. The item's overall probability of being answered correctly and its point bi-serial correlation with the total test score were also indicated.

If the mean probability of a correct response to a posttest item was found to be particularly low (less than .75), the item was examined closely for construction deficiencies. Specifically, the following four questions were asked regarding the item: (1) Is the stem meaningful? (2) Does the stem present a definite problem? (3) Does the stem contain irrelevant material? (4) Is the correct answer the only correct response or clearly the best response? In this case, even more reliance was placed on the pattern of responses to the distractor alternatives. If the item contained no obvious deficiencies, the instruction itself was suspect. The relative response patterns of students falling into the different pantiles were then examined. If the item tended to discriminate students in the lowest pantile (the bottom 20% with respect to total score) from students in the upper four pantiles, the instructional problem was not considered to be too serious. If, on the other hand, the pattern of incorrect responding was relatively even across all levels of overall student performance, a clear case was made for altering the instruction. Modification of the instruction was also considered if the mean correct response probability for all items pertaining to a particular objective was relatively low.

The third source of information for revision consisted of student statements recorded on the attitude questionnaires. If a relatively large percentage of students indicated that they required more than four hours to complete their study of the module, the amount of reading material assigned was reduced. This was done without eliminating objectives whenever possible. Students were also requested to indicate which of the readings in each module they would prefer to have eliminated. Alternative readings were sought to replace those readings which a particularly large proportion of students nominated for elimination.

Only minor changes were made in the instructional materials for the Computers in Education module. Several paragraphs were rewritten for clarity, one of the readings was replaced with a more interesting paper, and the on-line demonstration of computer-assisted instruction was deleted. Approximately one-quarter of the test items were replaced or modified.

Similarly, the revision of the Classroom Management module was relatively minor. Some definitions in the supplementary materials were rewritten for clarity and approximately one-fifth of the test items were revised.

Portions of several readings in the Tests and Measurement module were deleted, the series of readings was reorganized, and one paper was added for the purpose of providing an organizational overview. In all, the length of the module was reduced by one-quarter. In conjunction with this reorganization, several of the objectives were rewritten and the number of objectives reduced from 12 to 10. The number of test items on each form was reduced from 30 to 20. Of these 20, approximately one-quarter were either new items or items which had been modified.

The Cultural Differences module presented the major revision problem. It will be recalled that the module was much too time-consuming and that student performance on this module was substantially below that of the other four modules. The total length of the module was shortened by one-half. This was accomplished by deleting one reading and replacing two others by brief papers written by the module authors. Four of the 12 objectives were deleted and two others rewritten to reflect these changes. The number of test items per form was correspondingly reduced from 24 to 15. Of these 15 items, approximately two-thirds were either revised or new items.

It will be recalled that the Statistics module was divided into five units with a separate posttest for each unit. This was found to be a relatively unsatisfactory situation. Too much computer time was required to administer the tests and students objected to the requirement that they make repeated trips to the computer terminal room. When faced with a long series of tests, students who failed an initial posttest tended to retake the test immediately

rather than spend any substantial amount of time studying and returning to take the retest at a later date. Consequently, the module was reorganized into two rather than five units. Based on student performance, student comments on the attitude questionnaire and feedback from the course instructors, the material previously contained in one of the five units was deleted from the module. The number of objectives was correspondingly reduced from 16 to 13. A short supplementary reading was added to clarify a concept not adequately covered in the original materials The total number of test items for the module was reduced from 60 to 37 per form of which approximately one-third were new or modified.

The introduction to the set of modules was revised and expanded to provide students with more complete information regarding the procedures to be followed in the CMI portion of the curriculum. Finally, a five-page section was added to the manual which provided the students with a brief tutorial on the nature of criterion-referenced tests and the appropriate use of objectives as a guide to studying.

The revision of the five modules and their tests was completed by January 15, 1973. They are currently being used by approximately 450 student registered in the course during the spring semester. A second evaluation of student performance and attitudes will be conducted following their completion of the modules. It is anticipated that the revisions described above will result in an increase in posttest performance levels.

REFERENCES

Bloom, B. S. (Ed.), Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. Taxonomy of educational objectives. Handbook I: Cognitive domain. New York: McKay, 1956.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

Edmonston, L. P., Randall, R. S., & Oakland, T. A model for estimating the reliability and validity for criterion referenced measures. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1972.

Educational Psychology Modules, Computer-Assisted Instruction Laboratory and Department of Educational Psychology, The University of Texas at Austin, 1972.

Educational Testing Service. Tests and measurement kits. Berkeley, Calif.: ETS, 1971.

Gage, G. A model for establishing a priority of educational needs. Teaching Research, Monmouth, Oregon, 1970.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18, 519-21.

International Business Machines, Inc. IBM Coursewriter II Author's Guide, Part I: Course Planning. New York: IBM, 1967.

Jennings, E. ITEM A, A general purpose item analysis program. The University of Texas at Austin, 1970.

Kriewall, T. E. Aspects and applications of criterion referenced tests. Technical Paper 103. Institute for Educational Research, 1400 West Maple Avenue, Downers Grove, Illinois 60515. April 1972.

Leherissey, B. L. The validation of a measure of state epistemic curiosity in a computer-assisted learning situation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1972.

Mitchell, R. W., & Conner, N. H. User's guide to Coursewriter II preprocessor, V1.2. Systems Memo No. 3. Computer-Assisted Instruction Laboratory, The University of Texas at Austin, May 1971.

Richardson, F. C., O'Neil, H. F., Jr., Grant, R. D., Jr., & Judd, W. A. Development and preliminary evaluation of an automated test anxiety reduction program for a computer-based learning situation. Technical Report No. 20. Computer-Assisted Instruction Laboratory, The University of Texas at Austin, January 1973.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. Manual for the State-Trait Anxiety Inventory. Palo Alto, Calif.: Consulting Psychologist Press, 1970.

Smith, B. O., Stanley, W. O., & Shores, J. H. Fundamentals of curriculum development (Rev. Ed.). New York: Harcourt, Brace and World, 1957.

APPENDICES

APPENDIX A

COMPUTERS IN EDUCATION

COMPUTERS IN EDUCATION

Claire Weinstein
Department of Educational Psychology

The University of Texas at Austin
Summer, 1972

Instructional Materials:

I. Introduction to the Module
II. Demonstration of a Computer-Assisted Instruction Program (Sutton Hall 314)
III. Reading Assignments with Objectives and Sample Test Questions
IV. Six Mimeographed Readings

I. INTRODUCTION TO THE MODULE

This module is designed to provide the student with an introduction to one aspect of the radical changes taking place in educational thought and practice. As our conceptions of the school and its place in society change, so do the roles of the student and the teacher. The use of computers in the schools represents one of the areas of innovation destined to have a major impact on our total educational environment.

The materials you will read have been written or selected to introduce you to the principles underlying the design of programmed instructional materials (of which computer-assisted instruction is one version) and to provide you with an overview of a number of ways in which computers may be used in an educational setting. Two of the papers in the module discuss changes in the role of the teacher and the activities of students which can be expected when computer technology is introduced into a classroom. The final paper in the module is a brief assessment of the current positive aspects and shortcomings of instructional computer applications.

This module is not intended to sell the concept of computer applications to education. Hopefully, computers will be adopted for particular instructional applications on the merits of their usefulness in those specific situations. It is quite probable, however, that you will encounter some instructional computer application in your teaching career. This module is intended to provide you with enough information, negative as well as positive, to allow you to decide how you might use the computer's capabilities (if at all) in your own classroom.

As you work your way through this series of modules, you will be participating in one form of an instructional computer application--computer-managed instruction. Unfortunately, you, as a student, will not see much of how the system operates. Your only exposure to the computer will be while taking tests or obtaining prescriptive assignments. To give you a better understanding of the computer's instructional capabilities a demonstration program has been prepared as one component of this module. This program will teach you to use the computer terminal and includes a short segment of a highly interactive computer-assisted instruction program. Make an appointment to see this demonstration as soon as possible.

The complete appointment will take about one hour. We suggest that you get together with a friend and make a single appointment for the two of you. We have found that such demonstrations are much more interesting if two or more people work through them together.

It should take you about three hours to complete the readings in this module. Study the article description, objective(s), and sample test item(s) before reading each paper. The purpose of the objectives is to help you direct your study of the materials. When you have completed the module, you will be given a 15-item multiple-choice test. If you review the sample test items given with the objectives, they should give you a good indication of whether or not you are ready to take the mastery test.

II. DEMONSTRATION OF A COMPUTER-ASSISTED INSTRUCTION PROGRAM (Sutton Hall 314)

III. READING ASSIGNMENTS, OBJECTIVES, AND SAMPLE TEST ITEMS

After completing each of the reading assignments, you should have achieved the objective(s) given for that assignment. A sample test item is provided below each objective to allow you to assess your mastery of that objective. The correct answer is starred (*). The mastery test you will take at the computer terminal upon completing these readings is composed of items similar to the sample items.

Reading 1: Computer-Assisted Instruction (CAI) by Wilson A. Judd

This reading serves as an introduction to the use of computers in education by discussing what many consider to be their most interesting educational application--computer-assisted instruction or CAI. CAI is a form of programmed instruction. To provide a context in which to view CAI, instructional programming is discussed as a process for the development of instructional materials. Following a brief description of two earlier forms of programmed instruction, the general characteristics of CAI are discussed. Three specific types of CAI programs are then described: drill and practice; tutorial; and dialogue. Some of the assumptions underlying these types of programs are examined and a fairly recent innovation, learner-controlled CAI, is described.

Objective 1

The student will identify characteristics of instructional materials design for programmed instruction and CAI.

Sample Test Item

Which of the following is not a characteristic of instructional materials developed for CAI?

(1) The materials are auto-instructional.
(2) The materials are individualized.
*(3) All revisions are completed before the materials are administered.
(4) The materials provide the student with feedback on his performance.

Objective 2

The student will identify the definition of CAI and characteristics and examples of the following three types of CAI programs: (1) drill and practice; (2) tutorial; and (3) dialogue.

Sample Test Item

A high school Spanish class uses the school's computer terminals to practice their vocabulary and spelling skills. They are most probably using a

(1) simulation program.
(2) tutorial program.
*(3) drill and practice program.
(4) dialogue program.

Objective 3

The student will identify characteristics and examples of learner-controlled (as opposed to conventional) computer-assisted instruction.

Sample Test Item

A drill and practice program in arithmetic is designed so that problems of a given type continue to be presented to the student until he indicates that he is ready for the next problem type. This is an example of

(1) computer-managed instruction
*(2) learner-controlled instruct
(3) simulation.
(4) computer-based evaluation.

Reading 2: The Use of Computers for Instructional Simulation

This reading consists of two parts: a general description of the ways in which computers can be used to simulate real-life situations for instructional purposes and an article (The Application of Computer Technology to the Improvement of Learning by Don D. Bushnell) which discusses several specific instructional applications of computer simulation. Particular emphasis is placed on the teaching of disadvantaged children.

Objective 4

The student will identify the reasons for computer simulation and the benefits that simulation offers for disadvantaged children.

Sample Test Item

Which of the following is not a circumstance in which computer simulations are appropriate substitutes for real-life experiments?

(1) When the time scale causes difficulty.
(2) When danger is involved.
(3) When equipment is unavailable because of expenses or complexity.
*(4) When a qualified teacher is not available.

Sample Test Item

Mrs. Jones' school has just installed a computer-managed instructional system. As a teacher, her role will now be that of a

(1) lecturer.
(2) clerk.
(3) disseminator of information.
*(4) manager of the learning process.

Reading 6: General Assessment of Computer-Assisted Instructional Systems by Wilson A. Judd

This final reading is presented for the purpose of attempting to place the current and future educational roles of computers in their proper perspective. The computer is potentially a very powerful instructional medium. Currently, there are relatively few applications in which the computer's use has begun to approach this potential. Many technical, pedagogical and social problems remain to be solved. There will be no test questions on this reading.

APPENDIX B
CLASSROOM MANAGEMENT

CLASSROOM MANAGEMENT

W. Paul Scott
Department of Educational Psychology

and

Donald D. Rogers
Department of Curriculum and Instruction

The University of Texas at Austin
Summer, 1972

Instructional Materials:

I. Introduction to the Module
II. Cognitive Objectives and Sample Test Questions
III. Supplementary Materials
IV. Mink, O. G., The Behavior Change Process. New York: Harper & Row, 1970. Paperback, $4.95. Available at the Co-op and Hemphill's under your instructor's name.

I. INTRODUCTION TO THE MODULE

This module is designed to give you a new vocabulary to use in discussing classroom discipline problems and some insights about the way teachers and students influence each other's behavior. You will be introduced to some specific techniques for changing undesirable student behavior and replacing it with more appropriate behavior.

The materials you will read are a relatively short paperback by Oscar Mink, The Behavior Change Process, and some supplementary materials prepared here at the University of Texas. The book begins with two short pretests. You may find it interesting to take these tests to find out how much you already know about the material. The text itself is divided into two parts. A brief introductory section introduces you to the way a "behavioral" psychologist looks at student-teacher interactions and discipline problems and presents the basic terminology you'll be learning. The major segment of the book is a programmed course on the various techniques of changing student behavior. Since the material is presented in one unit, rather than in topical chapters, you will probably find it helpful to read the programmed portion of the book in as few sittings as possible. Our experience thus far indicates that 85% of the students complete Mink's book in four hours or less.

To assist you in your study, we have prepared a short set of supplementary materials which define and discuss some terms which we consider to be particularly critical to your understanding of this topic. The definitions of the terms have been indented, so that you can find them quickly.

After you've read the material, you will be given a 20-item multiple-choice mastery test. As is true for all the modules you'll use this semester, the items on the mastery test were written to measure your mastery of a set of cognitive objectives. The cognitive objectives for this module are listed on the following pages. Each is a statement of a specific learning goal for these materials. We have included a sample test item corresponding to each objective.

We suggest that you begin by studying the cognitive objectives and corresponding sample test items. Next, read over the supplementary materials to acquaint yourself with the new terminology. Then study the textbook. Finally, study the supplementary materials again to be sure that you understand the terms defined there. It is a good idea to use the posttests at the end of the textbook and to review the sample test items given below to determine whether or not you are ready to take the mastery test.

II. COGNITIVE OBJECTIVES AND SAMPLE TEST QUESTIONS

After completing this module, you will have achieved the following objectives. Notice that below each objective we have included a sample test question designed to assess your mastery of that objective. The correct answer is starred (*). The mastery test you take upon completing your study of Mink's programmed text will be composed of items similar to these.

Objective 1

When given the name of any term or process defined in the supplementary materials, the student will select the correct paraphrased definition of that term or process.

Sample Test Item 1

A positive reinforcer

(a) is the addition of an aversive stimulus.
(b) decreases the frequency of undesirable behavior.
*(c) increases the frequency of the behavior it follows.
(d) is the opposite of a punisher.

Objective 2

When given a description of any of the processes defined in the supplementary materials, the student will correctly identify and label that process.

Sample Test Item 2

Applying an aversive stimulus is an example of

(a) extinction.
*(b) punishment.
(c) shaping.
(d) negative reinforcement.

Objective 3

When given a list of descriptive phrases, the student will correctly select those phrases which describe observable behavior.

Sample Test Item 3

Which of the following is a behavior?

(a) thinking
(b) understanding
(c) knowing
*(d) touching

Objective 4

When given a description of a student-teacher interaction in a specific classroom situation or a general history of student-teacher interactions, the student will select the most likely future teacher and student behaviors.

Sample Test Item 4a

If, by asking silly questions, a student continually gets the teacher's attention, the student will probably

(a) stop asking questions.
*(b) continue to ask silly questions.
(c) try to ask an intelligent question.
(d) find other ways to get attention too.

Sample Test Item 4b

The teacher decided to ignore every silly question. Most students will probably

(a) stop asking silly questions immediately.
(b) ask more intelligent questions.
(c) continue to ask silly questions.
*(d) begin to ask silly questions less frequently.

Objective 5

When given a description of a specific class-teacher or child-teacher interaction, the student will correctly identify and label the components of that interaction.

Sample Test Item 5

In the situation above (sample test items 4a and 4b), the teacher learned that her attention

(a) negatively reinforced the desired response.
(b) negatively reinforced the undesired response.
(c) positively reinforced the desired response.
*(d) positively reinforced the undesired response.

Objective 6

When given a description of the class-teacher or child-teacher interaction, the student will correctly select the most appropriate teacher responses for producing a desirable change in student behavior.

Sample Test Item 6

In the situation above (sample test items 4a and 4b), in order to increase appropriate questions the teacher should

(a) continue to extinguish silly questions by ignoring them.
(b) punish silly questions every time they occur.
(c) counter condition new verbal responses.
*(d) shape more relevant questions by continually attending to them.

APPENDIX C
TESTS AND MEASUREMENT

TESTS AND MEASUREMENT IN THE CLASSROOM

Paul Dixon and Claire Weinstein
Department of Educational Psychology

The University of Texas at Austin
Summer, 1972

Instructional Materials:

I. Introduction to the Module
II. Reading Assignments with Objectives and Sample Test Questions
III. Five Mimeographed Readings Included with the Module
IV. Two booklets from the Educational Testing Service (ETS) Tests and Measurement Kit: (1) Selecting an Achievement Test, and (2) Making the Classroom Test

I. INTRODUCTION TO THE MODULE

This module is designed to familiarize you with a variety of evaluation procedures and instruments for use in the classroom. The characteristics and educational uses of standardized aptitude and achievement tests are discussed. A number of technical labels and terms are presented to facilitate your understanding and use of such tests. The various purposes of testing are discussed and, finally, a set of general and specific guidelines for developing your own classroom tests are presented.

The materials you will read consist of five short papers (included with the module) and two booklets from the Educational Testing Service (ETS) Tests and Measurement Kit. ETS is a nonprofit organization, located in Princeton, New Jersey, devoted to measurement and research in education. You will undoubtedly come into frequent contact with ETS products and services throughout your teaching career. There is a wide variety of other materials included in the kit which will probably be of interest to you now and in your future career. ETS's various testing programs and services are described together with information about several specific tests. Two booklets, one on preparing multiple-choice questions and another on short-cut statistics for classroom tests, should be of particular interest. Only two of the booklets, however, Selecting an Achievement Test and Making the Classroom Test, have been assigned as readings for this module.

Of a previous group of students studying a slightly longer version of this module, 81% indicated that they completed the readings in four hours or less. After you have read the material, you will be given a 30-item multiple-choice mastery test. If you review the sample test items given with the objectives after you have studied the materials, they should give you a good indication of whether or not you are ready to take the mastery test.

II. READING ASSIGNMENTS, OBJECTIVES, AND SAMPLE TEST QUESTIONS

After completing each of the reading assignments, you should have achieved the objectives given for each assignment. We have provided a sample test question below each objective designed to assess your mastery on that objective. The correct answer is starred (*). The mastery test you will take upon completing the readings will be composed of items similar to these.

Reading 1: Gathering Information About Students by C. A. Cartwright (adapted for CMI format).

The purpose of this reading is to provide the student with an introductory overview of some basic ideas about data collection in the classroom.

Objective 1

The student will recognize the characteristics of quantitative and qualitative information by identifying examples of each and by recognizing the rationale for reporting observed behavior.

Sample Test Item

Qualitative information is

*(1) not as precise as quantitative information.
(2) numerical information.
(3) obtained through objective testing.
(4) used only when we can precisely measure a specific ability.

Objective 2

The student will recognize the various steps in the teaching-learning process by identifying examples of and the characteristics of each step in the process.

Sample Test Item

Which of the following is not a step in the teaching-learning process?

(1) Provide feedback to the learner.
(2) Select and use teaching procedures.
(3) Formulate objectives.
*(4) Assign student grades.

Reading 2: Evaluation Procedures by C. A. Cartwright (adapted for CMI format).

This article provides a brief overview of the types of standardized tests used in the classroom.

Objective 3

The student will recognize the characteristics of aptitude and achievement tests and identify examples of each.

Sample Test Item

Aptitude tests are

*(1) designed to measure a student's capacity or potential.
(2) designed to measure accomplishment in a subject area.
(3) seldom standardized.
(4) usually administered individually by the teacher.

Reading 3: Reliability, Validity, and Usability by D. Schreiber and C. A. Cartwright (adapted for CMI format).

This reading describes various considerations which are important in selecting an existing testing instrument or the improvement of teacher-made instruments.

Objective 4

Given examples of the types of reliability and validity, the student will recognize characteristics of each, and will identify the relationship between reliability and validity.

Sample Test Item

A test that yields similar scores upon repeated administration to the same individuals is said to have high

(1) validity.
*(2) reliability.
(3) causality
(4) none of these

Reading 4: Selecting an Achievement Test: Principles and Procedures, Educational Testing Service. (The reader may exclude the sections on validity and reliability, pp. 7-11.)

This booklet provides the student with a number of guidelines for the proper selection of standardized tests for classroom use. The sections on reliability and validity in this article will not be tested. They are, however, recommended as an aid to understanding the material in Reading 3. Moreover, no objectives require computational procedures.

Objective 5

Given examples of groups to be tested, the student will select the proper norm group.

Objective 6

The student will recognize the various considerations necessary for the selection of the proper achievement test.

Objective 7

The student will recognize the purposes of testing and identify the characteristics of each purpose.

Sample Test Item

Which of the following is not a use of tests?

(1) placement analysis.
(2) diagnosis analysis.
(3) assessment analysis.
*(4) subjective analysis.

Reading 5: Interpreting Test Scores Realistically, author unknown.

This reading offers a number of important considerations necessary for correct interpretation of test scores from both standardized and teacher-made tests. A test score is only an estimate of the student's ability at any given time and it takes on meaning only when it is properly interpreted.

Objective 8

The student will interpret the relationship between two pupils' test scores when given (a) the two scores and (b) the standard error of measurement for the test.

Objective 9

The student will recognize the 'guiding principles' of test interpretation.

Sample Test Item

When interpreting tests, the teacher must always remember that

(1) a student's score alone indicates the grade he should receive.
(2) a student's score indicates his true ability.
(3) "true score" is another term for the scores obtained on any given test.
*(4) percentiles gain meaning only when the characteristics of the norm group are known.

Reading 6: Teacher Devised Achievement Tests by C. A. Cartwright (adapted for CMI format).

This section offers a general overview of the various methods used in the construction of teacher-made tests. Several question formats are discussed.

Objective 10

The student will recognize the specific characteristics and special uses of each of the question formats discussed.

Sample Test Item

If we wished to measure recall of information that has been stored in memory, we would use

*(1) completion items.
(2) multiple-choice items.
(3) alternative response items.
(4) matching exercises.

Reading 7: Making the Classroom Test: A Guide for Teachers, Educational Testing Service.

This booklet is a practical guide to the construction of better test questions. Don't lose it! It will be a very helpful aid when you are actually confronted with your first test construction task.

Objective 11

When given an example test question, the student will recognize three obvious pitfalls that should be avoided in writing test questions: specific determiners; extreme reading difficulty; and ambiguity.

Objective 12

The student will recognize the limitations and advantages of both essay and objective test questions.

Sample Test Item

An objective test item

(1) cannot tap high levels of reasoning.
(2) covers a narrow field of knowledge.
(3) can be scored accurately and consistently.
*(4) all of the above.

APPENDIX D
CULTURAL DIFFERENCES

CULTURAL DIFFERENCES

Brenda Rutherford and Charles D. Saddler

Department of Educational Psychology
The University of Texas at Austin
Fall, 1972

Instructional Materials:

I. Introduction to the Module
II. Reading Assignments with Objectives and Sample Test Questions
III. Six Readings

I. INTRODUCTION TO THE MODULE

This module is designed to familiarize you with certain aspects of cultural differences between the Anglo middle-class culture and the black and Mexican-American cultures. The major areas of cultural differences covered in this module are language and standardized test performance differences and how these differences affect school achievement.

The materials you will read consist of six articles drawn from a variety of sources. The first two articles deal with differences in language between the black and Mexican-American cultures and the Anglo middle-class culture. The following two articles discuss IQ and standardized test performance differences on the part of blacks and Mexican-Americans. The fifth article summarizes cultural differences in language, IQ, and test performance, and concludes with a discussion of how these differences are relevant to the present educational situation. The final article presents a formal discussion of cultural differences in intellect and how these differences affect school performance.

It is anticipated that you should complete the readings in four hours or less. After you have read the material, you will be given a multiple-choice mastery test. If you review the sample test items given with the objectives after you have studied the materials, they should give you a good indication of whether or not you are ready to take the mastery test.

II. READING ASSIGNMENTS, OBJECTIVES, AND SAMPLE TEST QUESTIONS

After completing each of the reading assignments, you should have achieved the objectives given for each assignment. The sample test items provided below each objective are designed to assess your mastery of that objective. The correct answer is starred (*). The mastery test you will take upon completing the readings will be composed of items similar to these.

Reading 1: Academic Ignorance and Black Intelligence, by William Labov.

Labov discusses the basic tenets of the verbal deprivation theory and then presents evidence to refute the theory. He demonstrates that, although black children appear to be non-verbal in certain situations, they are quite verbal in a familiar environment when using the Black English Vernacular (BEV). Labov presents an interesting discussion of the merits of middle and lower class speech, questioning the dominance of middle class values in education. He concludes that the failure of education has been a result, not of the verbal deprivation of the black children, but of the inability of the school to adjust and adapt to cultural differences.

Objective 1

The student will identify characteristics of the verbal deprivation theory as it relates to Black English.

Sample Test Item:

According to the verbal deprivation theory, black children in the urban ghetto areas

1. participate fully in a highly verbal culture.
2. seldom participate in verbal interchanges with peers.
*3. cannot form concepts or convey logical thoughts.
4. have the same basic vocabulary as white children.

Objective 2

The student will identify characteristics of Labov's linguistic refutation of the verbal deprivation theory and recognize the implications for teaching.

Sample Test Item:

According to Labov, one serious consequence of the verbal deprivation theory is its negative effect on

*1. the teacher's attitude toward the student.
2. interpersonal relationships among students.
3. the attitude of the student toward the teacher.
4. the types of interviews which are successful.

Reading 2: Bilingually Advantaged, by Judith Rae Gates.

Gates presents four viewpoints prevalent in approaches to educating Mexican-American children. She mentions the bilingual approach and gives justification as to why it is better than any of the other approaches. The remainder of the article discusses a bilingual program currently being implemented in San Antonio, Texas.

Objective 3

The student will identify differences between the English as a Second Language (ESL) approach and the bilingual approach for the education of Mexican-American children.

Sample Test Item:

The bilingual approach for the education of Mexican-American children differs from the English as a Second Language (ESL) approach in that

1. the bilingual approach is based on methods used in teaching a foreign language.
*2. the ESL approach is based on methods used in teaching a foreign language.
3. the ESL approach requires that 400 English words be learned at the beginning.
4. the bilingual approach initially makes no use of English for classroom interaction.

Objective 4

The student will recognize arguments for the bilingual approach as presented by the author and as demonstrated in the San Antonio system.

Sample Test Item:

The bilingual approach as implemented in the San Antonio school system is advantageous because it

1. teaches the student 400 words in English in the first grade.
2. groups students according to their knowledge of English.
*3. makes the curriculum more relevant to Mexican-American experiences.
4. tests the students to evaluate their knowledge of English.

Reading 3: The I.Q., Chapter 2 in Herbert Ginsburg's The Myth of the Deprived Child.

Ginsburg discusses four major "myths" concerning intelligence (IQ) tests. He illustrates why these myths are not valid. Ginsburg also discusses the controversy concerning the effects of environment and heredity on I.Q.

Objective 5

The student will recognize the shortcomings of measured "intelligence" (defined by IQ tests) as described by Ginsburg.

Sample Test Item:

According to Ginsburg, an important quality which IQ tests do not measure is

1. verbal skills.
*2. creativity.
3. memory.
4. perception.

Objective 6

The student will identify characteristics of heredity and environment that may affect performance on IQ tests.

Sample Test Item:

Ginsburg presents evidence to show that one of the most important factors affecting IQ scores is

1. age.
2. sex.
3. number of siblings.
*4. socio-economic status.

Reading 4: Assessing Assessment Instruments: A Chicano Perspective, by Ernest M. Bernal, Jr.

Bernal discusses the testing situation as it currently applies to Mexican-Americans. He lists six ways in which the use of tests may discriminate against Mexican-Americans. However, he believes that testing can serve a legitimate purpose if used for understanding and planning, rather than merely labeling. He concludes by recommending 10 courses of action to be taken in order to end the discriminating use of tests with Mexican-Americans. (It is important to realize that, although Bernal confines his discussion to test discrimination against Mexican-Americans, what he says can be applied to many minority groups.)

Objective 7

The student will identify ways in which tests are inappropriately used with members of minority groups.

Sample Test Item:

Which of the following is a way in which tests are inappropriately used with members of minority groups?

1. Using tests for program modification.
2. Using tests for diagnostic purposes.
*3. Using tests that are irrelevant to minority group experience.
4. Using tests that are criterion-referenced.

Objective 8

The student will identify (a) reasons why tests designed for Anglos should not simply be renormed and used with Mexican-Americans, and (b) Bernal's recommendations for improving the testing of Mexican-American students.

Sample Test Item:

Which of the following does Bernal recommend in order to improve the testing of Mexican-American children?

1. Test Mexican-American children only after lunchtime.
2. Use only Mexican-American testers to test Mexican-American children.
3. Use only tests which are written in "standard" Spanish.
4. Develop new testing instruments and procedures to minimize the gap between ethnic groups.

Reading 5: Cultural Differences and Inferences about Psychological Processes, by Michael Cole and Jerome S. Bruner.

Cole and Bruner discuss two interpretations of ethnic and social class differences: the deficit interpretation, which postulates that poor children are deficient because of their environment; and the difference interpretation, which postulates that poor children are not deficient, but different in that they demonstrate skills which are different from traditional middle-class skills. Evidence is presented which supports the latter interpretation. Cole and Bruner relate this evidence to the concepts of competence and performance. The remainder of the article is devoted to a discussion of the relationship between these two concepts, and the implications for teaching.

Objective 9

The student will identify characteristics of the concepts of competence and performance, especially as they relate to "cultural deprivation."

Sample Test Item:

Which of the following is true regarding the concepts of competence and performance discussed by Cole and Bruner?

1. Culturally deprived people have no competence.
*2. Competence is inferred from performance.
3. Performance is inferred from competence.
4. Performance is a hypothetical construct.

Objective 10

The student will identify implications from Cole and Bruner's presentation for teachers of "disadvantaged" children.

Sample Test Item:

An implication of the arguments set forth by Cole and Bruner is that the teacher of "disadvantaged" children should

1. tolerate classroom disorder.
2. use only traditional materials in the classroom.
3. maintain order at all costs.
*4. use relevant study materials.

Reading 6: Intellect and the Schools, Chapter 6 in Herbert Ginsburg's The Myth of the Deprived Child.

Ginsburg lists the assumptions underlying the government subsidized compensatory education programs currently in operation and examines the fallacies in each assumption. He then makes a strong case for new and innovative practices to be employed in the education of poor children.

Objective 11

The student will recognize the assumptions underlying compensatory education and Ginsburg's refutation of these assumptions.

Sample Test Item:

Compensatory education was created under the assumption that it would be

1. a method of making education more equitable.
2. the answer to the present educational crisis in urban schools.
*3. a means of enhancing school performance in poor children.
4. a method of providing day care services for poor working mothers.

Objective 12

The student will identify the characteristics of traditional education which, in Ginsburg's opinion, must be changed to make education beneficial to poor children.

Sample Test Item:

According to Ginsburg, traditional educational practices include

*1. a relatively small amount of freedom for the student.
2. open classroom situations for the student.
3. use of individualized instruction.
4. purple water fountains.

APPENDIX E
STATISTICS

STATISTICS IN THE CLASSROOM*

Paul N. Dixon
Department of Educational Psychology
The University of Texas at Austin

Fall, 1972

Instructional Materials:

I. Introduction to the Module.
II. Reading assignments with objectives and sample test questions.
III. Hereford, C. F., Natalicio, L. F. S., and McFarland, S. J., Statistics and Measurement in the Classroom. Kendall/Hunt, 1969. Paperback, $3.75. Available at the Co-op under your instructor's name.
IV. "Statistics: Tools for Better Teaching," a mimeographed paper by Paul N. Dixon, included with the module.
V. A Glossary of statistical terms and symbols, included with the module.

INTRODUCTION TO THE MODULE

The purpose of this module is to provide you with a number of statistical methods useful in the organization and interpretation of data from teacher-made tests and standardized tests.

The materials you will study consist of selected sections from a book of readings, Statistics and Measurement in the Classroom, a mimeographed paper included with the module, and a Glossary of statistical terms and symbols. This module differs from the others in this program in that it is divided into five units. Each unit has its own set of reading assignments and objectives. Also, rather than having one long test at the end of the module, there is a short multiple-choice test for each of the five units. The lengths of the tests are as follows: Unit I - 6 items, Unit II - 6 items, Unit III - 16 items, Unit IV - 20 items, and Unit V - 12 items. The testing procedure for this module differs from that of the other modules in that if you fail to reach criterion on the first test, you will be retested with an alternate form of the complete test rather than being retested over just the individual objectives failed on the first test. Since some of the tests are quite short, you might wish to take more than one test during a single session. You will be able to continue to a second test only if you reach criterion on the first test. If not, you will have to make another appointment to take the retest before continuing to the next unit test.

Of the group of 25 students who have previously studied this module, 71% indicated that they completed the readings in four hours or less.

READING ASSIGNMENTS, OBJECTIVES, AND SAMPLE TEST QUESTIONS

After completing each of the reading assignments, you should have achieved the objectives given for that assignment. A sample test question

*Used by permission of the author; copyright pending.

is provided for each of the units designed to assess your mastery of that unit. The correct answer is starred (*). The mastery test you will take after completing each unit is composed of items similar to these.

UNIT I

USES OF STATISTICS AND INTRODUCTION TO TERMS

Reading 1: "Statistics: Tools for Better Teaching," by Paul Dixon and page ix, Statistics and Measurement in the Classroom.

These readings explain why the teacher should study statistics and they describe in general terms the use of statistics in the classroom.

Objective 1

The student will identify the basic uses of statistics in the classroom.

Objective 2

The student will recognize the characteristics of evaluation discussed in the readings, the kinds of evaluation and their respective roles in the classroom.

Reading 2: "How to Talk Back to a Statistic," by Hull and Geis, Statistics and Measurement in the Classroom, pp. 59-74.

This article offers yet another reason why the teacher and most anyone else in our society should have some understanding of statistics. With the constant use of statistics in our news media, a basic understanding helps the teacher sort out distortions (through the almighty number) of many issues.

Objective 3

Given an experimental situation (e.g., sample selection), the student will recognize the validity or lack of validity of the statistical procedure.

Reading 3: "Statistical Terms and Statements," by Allen Edwards, Statistics and Measurement in the Classroom, pp. 3-7.

This article provides expanded definitions of various statistical terms found in your glossary and later in the readings. The article should be used in conjunction with the glossary as you may wish to refer to it when further explanation of the following terms is needed.

1. variability
2. relationship
3. average
4. sample
5. prediction
6. population
7. statistic
8. parameter
9. statistical inference
10. confidence limits
11. confidence coefficient
12. tests of significance

There will be no testing over this article.

Sample Test Item for this Unit

Statistics is to assessment evaluation as

1. testing is to judgmental evaluation.
2. scores are to subjective analysis.
3. assessment evaluation is to judgmental evaluation.
*4. subjective analysis is to judgmental evaluation.

UNIT II

VARIABLES AND SCALES

Reading 1: "Variables and Scales," by Allen Edwards Statistics and Measurement in the Classroom, pp. 9-21.

This reading is designed to provide the reader with a definition of the term "variable" and a description of the various measurement scales and their application.

Objective 1

The student will identify examples of (a) organismic, behavioral, stimulus, and response-inferred organismic variables, (b) the four scales of measurement, (c) discrete and continuous variables.

Objective 2

The student will recognize characteristics of (a) the four scales of measurement, and (b) discrete and continuous variables.

Sample Test Item for this Unit

The measurement scale in which the zero point is fixed is the

1. nominal scale.
2. ordinal scale.
3. interval scale.
*4. ratio scale.

UNIT III

FREQUENCY DISTRIBUTIONS, MEASURES OF CENTRAL TENDENCY AND THE NORMAL CURVE

Reading 1: "Statistical Analysis of Data," by Arnold Lien, Statistics and Measurement in the Classroom, pp. 23-30 (read to part F, p. 30), and p. 41, summary statements 1-5.

This section acquaints the reader with the methods of grouping data (such as test scores) and determining the three measures of central tendency, mean, median, and mode.

Objective 1

Given a set of raw scores, the student will construct a simple frequency distribution and compute the mean, median, and mode for that distribution of scores.

Objective 2

Given a specific situation, the student will choose the measure of central tendency best used in those circumstances, recognize the characteristics of the three measures of central tendency--mean, median, and mode.

Reading 2: "The Concept of the Normal Curve, by Susan McFarland, Statistics and Measurement in the Classroom, pp. 43-47.

This article further explains methods of handling data in addition to a description of the theoretical normal curve.

Objective 3

Given a graph of scores, the student will identify and differentiate a frequency polygon and a histogram and recognize the characteristics of each.

Objective 4

The student will recognize four characteristics of the theoretical normal curve.

Sample Test Item for this Unit

Which of the following is a characteristic of the mean, median, and mode?

*1. They may be statistics or parameters.
2. They are all measures of variability.
3. They are not subject to sampling error.
4. None of the above.

UNIT IV

MEASURES OF VARIABILITY

Reading 1 "Statistical Analysis of Data, by Arnold Lien, Statistics and Measurement in the Classroom, pp. 30-36 (read part F only), and p. 41, summary statement 6.

This section describes three measures of variability and the methods for calculating each.

Objective 1

Given an incomplete calculations table, the student will complete the table (see table 9, p. 84, Statistics and Measurement in the Classroom).

Objective 2

The student will recognize the characteristics of the external range, range by quartiles, and standard deviation and given a set of data on a calculations table (see table 9, p. 84, Statistics and Measurement in the Classroom), and the necessary formulas, calculate each.

Reading 2: "Uses of the Standard Deviation," by Abraham Franzblau, Statistics and Measurement in the Classroom, pp. 49-57.

Having learned how to compute the standard deviation, this article explains why you go to the trouble!

Objective 3

The student will recognize the relationship of the standard deviation to the normal curve.

Objective 4

Given a specific testing situation, the student will choose the best interpretation of the results and identify aspects of the standard deviation useful to the interpretation of test scores.

Objective 5

Given test scores and the standard deviation for two groups, the student will rank placement on these tests.

Sample Test Item for this Unit

If we wish to compare John's score on Test A with his score on Test B, which of the following statistical tools would be most helpful in providing a meaningful picture of his performance?

1. Range and correlation coefficient for the two tests.
2. Correlation coefficient and standard deviation for the two tests.
3. The median, mode, and range for the two tests.
*4. The mean and standard deviation for the two tests.

UNIT V

CORRELATION

Reading 1: "Statistical Analysis of Data," by Arnold Lien, Statistics and Measurement in the Classroom, pp. 36-41 (part G only) and pp. 41-42, summary statements 7-10.

This section gives a basic description and definition of correlation and it explains computation for the ρ(rho) or rank-difference correlation method.

Objective 1

Given an uncompleted correlation table (see table 10, p 38. Statistics and Measurement in the Classroom) and the formula for a rank-difference correlation, the student will complete the steps in calculating the table and compute a rank-difference correlation coefficient.

Objective 2

The student will recognize varying degrees of relationship and the two basic kinds of correlation from given examples, identify unwarranted causality statements derived from given data, and recognize the various characteristics of correlation.

Sample Test Item for this Unit

The correlation coefficient

1. indicates central tendency.
2. indicates which of two factors causes the other.
3. indicates an inverse relationship when its value is zero.
*4. can indicate as much relationship in a negative direction as in a positive direction.

APPENDIX F

DATA MANAGEMENT SYSTEM
FILE STRUCTURES

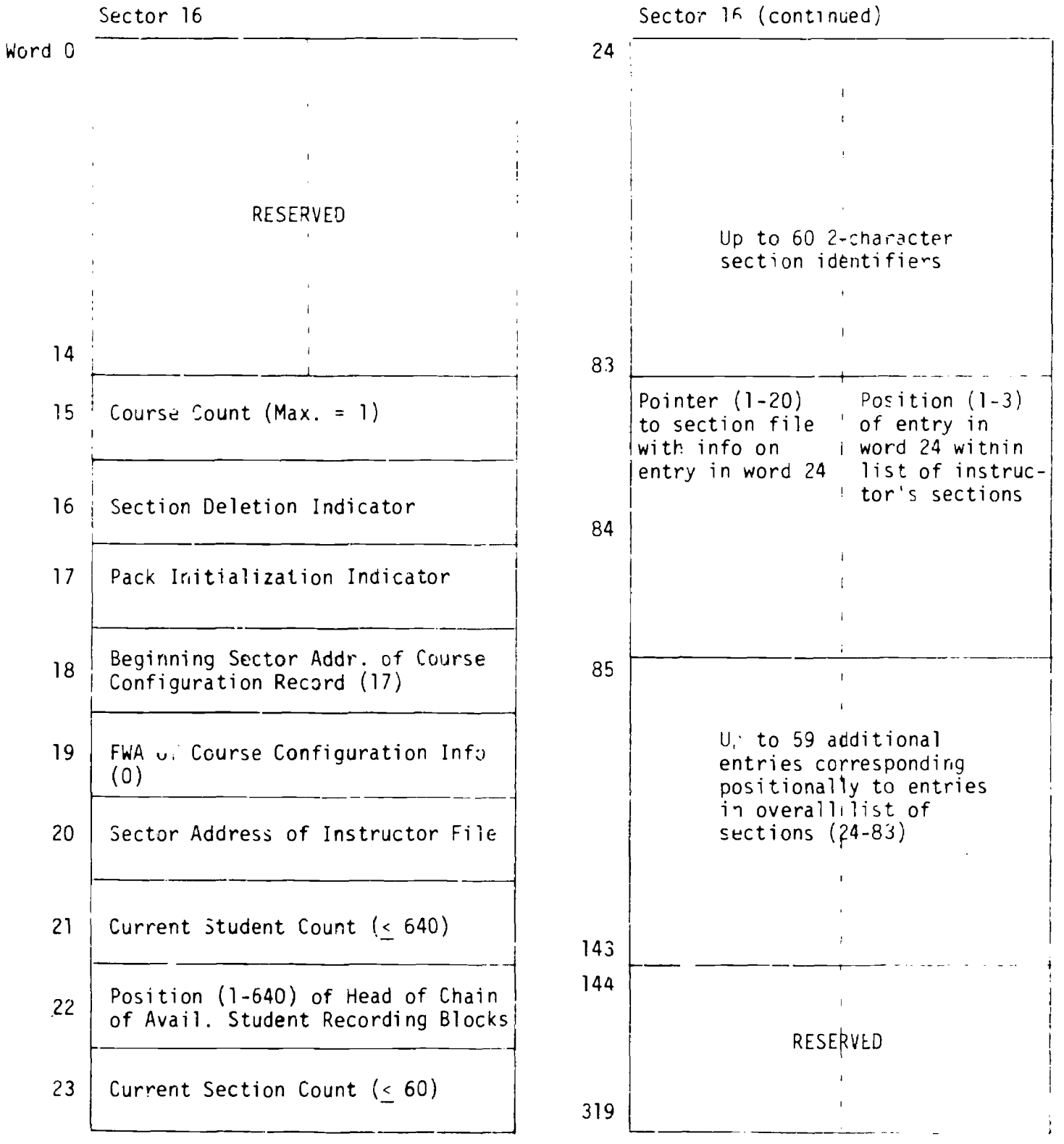


Figure 1. Structure of the recording pack directory (bookkeeping information)

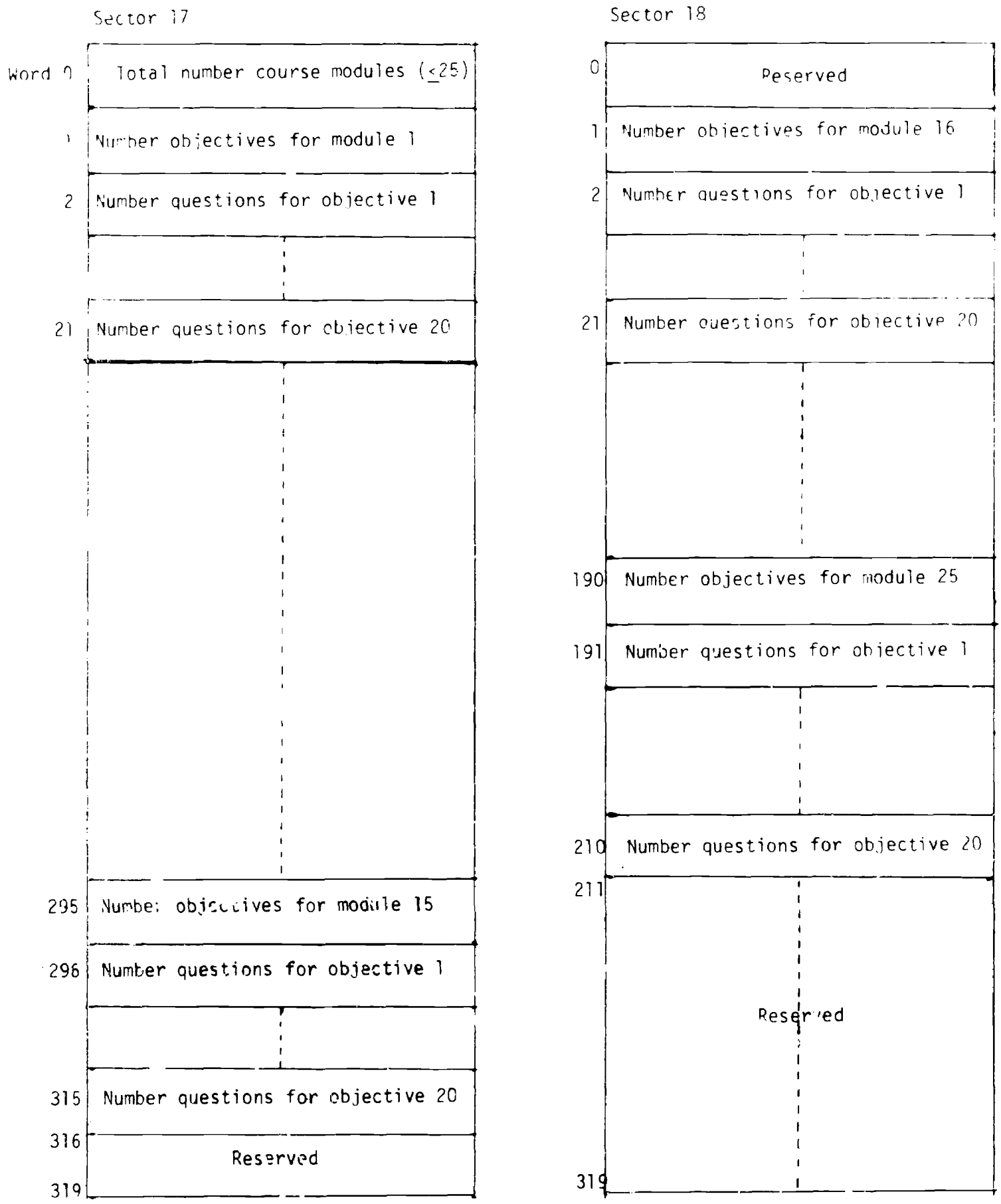


Figure 2. Structure of the course configuration record

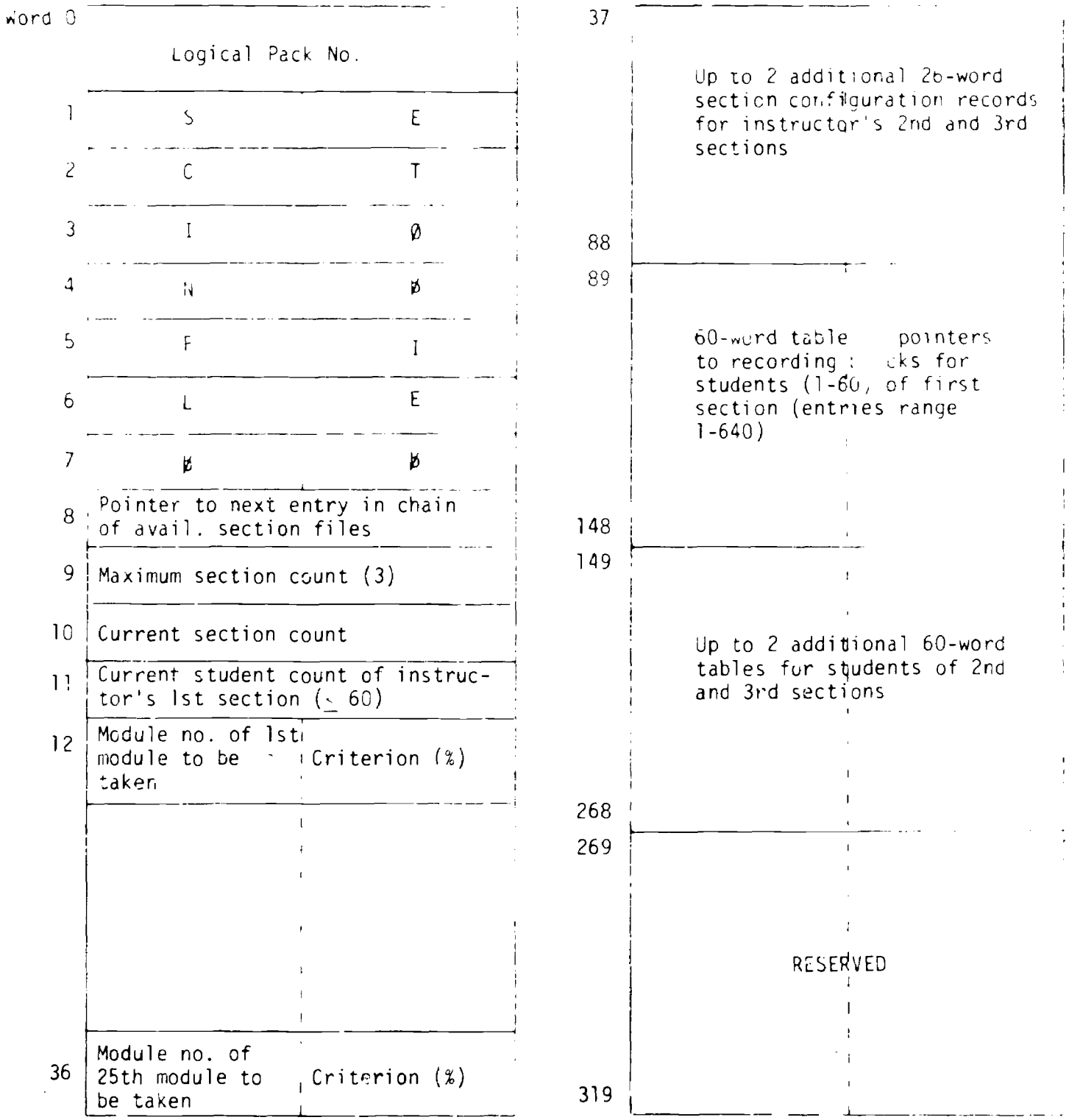
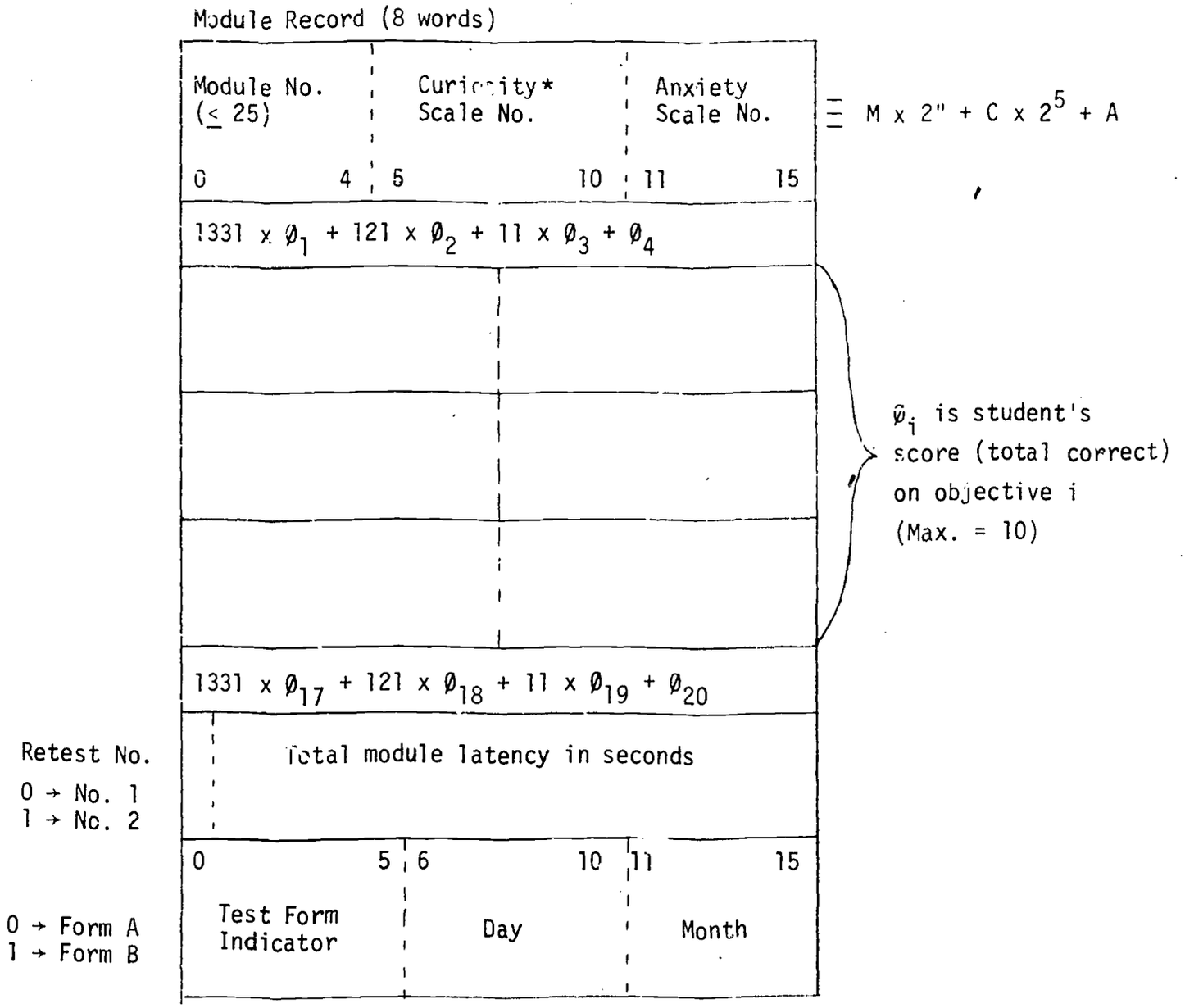


Figure 3. Structure of class section configuration information (disk sectors 22-41)



Sectors 300-1579 are designated as recording blocks. Each recording block consists of 2 sectors. Each student registered is allocated one block. Within that block, a series of 8-word records is written, one record for each module test/retest.

*Note: Test/retest is indicated by the value of the curiosity scale number. Test is indicated by a positive value. Retest is indicated by zero.

Figure 4. Structure of student data files (sectors 300-1579)